AJ

PCT WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 5 :		(11) International Publication Number:	WO 94/10203
C07K 15/00	A2	(43) International Publication Date:	11 May 1994 (11.05.94)

(21) International Application Number: PCT/US93/10520

(22) International Filing Date: 2 November 1993 (02.11.93)

(30) Priority data:

971,091	3 November 1992 (03.11.92)	US
029,335	4 March 1993 (04.03.93)	US
040,510	31 March 1993 (31.03.93)	US

(71) Applicant: CREATIVE BIOMOLECULES, INC. [US/US]; 45 South Street, Hopkinton, MA 07148 (US).

(72) Inventors: OPPERMANN, Hermann ; 25 Summer Hill Road, Medway, MA 02053 (US). OZKAYNAK, Engin ; 44 Purdue Drive, Milford, MA 01757 (US). KUBERA-SAMPATH, Thangavel ; Six Spring Street, Medway, MA 02053 (US). RUEGER, David, C. ; 19 Downey Street, Hopkinton, MA 01748 (US). PANG, Roy, H., L. ; 15 Partridge Road, Etna, NH 023750 (US). COHEN, Charles, M. ; One Harrington Lane, Weston, MA 02193 (US).

(74) Agent: KELLEY, Robin, D.; Testa, Hurwitz & Thibeault, Exchange Place, 53 State Street, Boston, MA 02109 (US).

(81) Designated States: AT, AU, BB, BG, BR, CA, CH, CZ, DE, DK, ES, FI, GB, HU, JP, KP, KR, LK, LU, MG, MN, MW, NL, NO, NZ, PL, PT, RO, RU, SD, SE, SK, UA, European patent (AT, BE, CH, DE, DK, ES, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG).

Published
Without international search report and to be republished upon receipt of that report.

(54) Title: OP-3-INDUCED MORPHOGENESIS

(57) Abstract

Disclosed are (1) nucleic acid and amino acid sequences for a novel morphogenic protein; (2) methods for producing and expressing the protein in a biologically active form; and (3) methods for utilizing the protein to induce tissue morphogenesis in a mammal, including methods for increasing a progenitor cell population in a mammal, methods for stimulating progenitor cells to differentiate and maintain their differentiated phenotype *in vivo* or *in vitro*, methods for inducing tissue-specific growth *in vivo* and methods for the replacement of diseased or damaged tissue *in vivo*.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AT	Austria	GB	United Kingdom	MR	Mauritania
AU	Australia	GE	Georgia	MW	Malawi
BB	Barbados	GN	Guinea	NE	Niger
BE	Belgium	GR	Greece	NL	Netherlands
BF	Burkina Faso	HU	Hungary	NO	Norway
BG	Bulgaria	IE	Ireland	NZ	New Zealand
BJ	Benin	IT	Italy	PL	Poland
BR	Brazil	JP	Japan	PT	Portugal
BY	Belarus	KE	Kenya	RO	Romania
CA	Canada	KG	Kyrgystan	RU	Russian Federation
CF	Central African Republic	KP	Democratic People's Republic of Korea	SD	Sudan
CG	Congo			SE	Sweden
CH	Switzerland	KR	Republic of Korea	SI	Slovenia
CI	Côte d'Ivoire	KZ	Kazakhstan	SK	Slovakia
CM	Cameroon	LI	Liechtenstein	SN	Senegal
CN	China	LK	Sri Lanka	TD	Chad
CS	Czechoslovakia	LU	Luxembourg	TG	Togo
CZ	Czech Republic	LV	Latvia	TJ	Tajikistan
DE	Germany	MC	Monaco	TT	Trinidad and Tobago
DK	Denmark	MD	Republic of Moldova	UA	Ukraine
ES	Spain	MG	Madagascar	US	United States of America
FI	Finland	ML	Mali	UZ	Uzbekistan
FR	France	MN	Mongolia	VN	Viet Nam
GA	Gabon				

OP-3-INDUCED MORPHOGENESIS

Field of the Invention

This invention relates generally to the field of tissue morphogenesis and more particularly to a novel protein that induces tissue morphogenesis in mammals.

Background of the Invention

Cell differentiation is the central characteristic of morphogenesis which initiates in the embryo, and continues to various degrees throughout the life of an organism in adult tissue repair and regeneration mechanisms. The degree of morphogenesis in adult tissue varies among different tissues and is related, among other things, to the degree of cell turnover in a given tissue. On this basis, tissues can be divided into three broad categories: (1) tissues with static cell populations such as nerve and skeletal muscle where there is no cell division and most of the cells formed during early development persist throughout adult life; (2) tissues containing conditionally renewing populations such as liver where there is generally little cell division but, in response to an appropriate stimulus, cells can divide to produce daughters of the same differentially defined type; and (3) tissues with permanently renewing populations including blood, testes and stratified squamous

epithelia which are characterized by rapid and continuous cell turnover in the adult. Here, the terminally differentiated cells have a relatively short life span and are replaced through proliferation of a distinct subpopulation of cells, known as stem or progenitor cells.

The cellular and molecular events which govern the stimulus for differentiation of these cells is an area of intensive research. In the medical field, it is anticipated that the discovery of factor(s) which control cell differentiation and tissue morphogenesis will advance significantly medicine's ability to repair and regenerate diseased or damaged mammalian tissues and organs. Particularly useful areas include reconstructive surgery and in the treatment of tissue degenerative diseases including arthritis, emphysema, osteoporosis, cardiomyopathy, cirrhosis, and degenerative nerve diseases.

A number of different factors have been isolated in recent years which appear to play a role in cell differentiation. Recently, various members of the structurally related proteins of the transforming growth factor (TGF)-β superfamily of proteins have been identified as true morphogens.

This "family" of proteins, sharing substantial amino acid sequence homology within their morphogenically active C-terminal domains, including a conserved six or seven cysteine skeleton, are capable

of inducing tissue-specific morphogenesis in a variety of organs and tissues, including bone, cartilage, liver, dentin, periodontal ligament, cementum, nerve tissue and the epithelial mucosa of the gastrointestinal tract. The proteins apparently bind to surface receptors or otherwise contact and interact with progenitor cells, predisposing or stimulating the cells to proliferate and differentiate in a morphogenically permissive environment. The morphogens are capable of inducing the developmental cascade of cellular and molecular events that culminate in the formation of new organ-specific tissue, including any vascularization, connective tissue formation, and nerve ennervation as required by the naturally occurring tissue.

Among the proteins useful in tissue morphogenesis are proteins originally identified as bone inductive proteins, such as the OP-1, (also referred to in related applications as "OP1"), OP-2 (also referred to in related applications as "OP2"), and the CBMP2 proteins, as well as amino acid sequence-related proteins such as BMP5, BMP6 and its murine homolog, Vgr-1, DPP and 60A (from Drosophila), Vgl (from Xenopus), and GDF-1 (from mouse) see, for example, U.S. Patent No. 5,011,691 to Oppermann et al., Lee (1991) PNAS 88: 4250-4254, and Wharton et al. (1991) PNAS 88: 9214-9218. These TGF-β superfamily members comprise a distinct subfamily of proteins different from other members of the TGF-β superfamily in that the family of morphogenic proteins are able to induce the full

cascade of events that result in tissue morphogenesis, including stimulating cell proliferation and cell differentiation of progenitor cells, and supporting the growth and maintenance of differentiated cells. The morphogenic proteins apparently can act as endocrine, paracrine or autocrine factors. Specifically, the endogenous morphogens may be synthesized by the cells on which they act, by neighboring cells, or by cells of a distant tissue, the secreted protein being transported to the cells to be acted on. In addition, the family of morphogenic proteins induce true tissue morphogenesis, rather than inducing formation of fibrotic (scar) tissue as, for example, TGF-β does.

The morphogens are synthesized in the cell as a precursor molecule approximately three times larger than the mature protein that is processed to yield mature disulfide-linked dimers comprising the C-terminal domain of the precursor sequence. The proteins are inactive when reduced e.g., in monomeric form, but are active as oxidized homodimeric species as well as when oxidized in combination with other morphogens under conditions to produce heterodimers. The proteins useful in tissue morphogenesis typically require a suitable environment enabling cells to migrate, proliferate and differentiate in a tissue-specific manner into, e.g., cartilage-producing chondroblasts, bone-producing osteoblasts, hemopoietic cells, or liver cells, depending on the nature of the

local environment. The proliferation and differentiation of cells induced by the morphogenic proteins requires a suitable local environment, including a suitable substratum on which the cells can anchor. The proliferating and differentiating cells also require the presence of appropriate signals to direct their tissue-specificity, such as cell surface markers.

It is an object of this invention to provide a novel purified morphogenic protein, "OP-3", including the amino acid sequence defining it and nucleic acids encoding it, including allelic, species, chimeric, and other amino acid sequence variants thereof, whether naturally occurring or biosynthetically constructed, and methods for utilizing the protein to induce the developmental cascade of tissue morphogenesis for a variety of tissues in mammals. The morphogenic properties of OP-3 include the ability to induce proliferation and differentiation of progenitor cells, and the ability to support and maintain the differentiated phenotype through the progression of events that results in the formation of adult tissue. Another object is to provide methods for the expression and isolation of morphogenically active species of OP-3 using recombinant DNA techniques. Yet another object is to provide generic sequences defining useful morphogens. Still another object is to provide tissue-specific acellular matrices that may be used in combination with OP-3, and methods for their

preparation. Other objects include utilizing OP-3 in a variety of applications including methods for increasing a progenitor cell population in a mammal; methods for stimulating progenitor cells to differentiate in vivo or in vitro and to maintain their differentiated phenotype; methods for inducing tissue-specific growth in vivo, and methods for the replacement of diseased or damaged tissue in vivo. These and other objects and features of the invention will be apparent from the description, drawings, and claims which follow.

Summary of the Invention

A novel substantially pure genetic sequence encoding a novel substantially pure protein referred to herein as "OP-3" now has been discovered. This novel protein is a member of the morphogenic protein family previously described by Applicants (see, for example, US92/01968 (WO92/15323), and US92/07432 (WO93/05751). Accordingly, the invention provides methods for utilizing OP-3 to induce the developmental cascade of tissue morphogenesis in a mammal. Specifically, methods are provided for utilizing OP-3 to induce the proliferation of uncommitted progenitor cells, to induce the differentiation of these stimulated progenitor cells in a tissue-specific manner under appropriate environmental conditions, and to support the growth and maintenance of these differentiated cells. The protein also may be used to stimulate the "redifferentiation" of cells that have strayed from their differentiated phenotypes. Accordingly, OP-3 can be utilized to initiate and maintain the developmental cascade of tissue morphogenesis in an appropriate, morphogenically permissive environment.

As used herein, useful OP-3 morphogens include proteins encoded by the DNA sequence provided in Seq. ID No. 1 ("mOP-3") and allelic and species variants thereof, as well as other naturally-occurring and biosynthetic amino acid sequence variants, including chimeric proteins, that are morphogenically active as

defined herein. "Morphogenically active fragment" is understood to include all proteins and protein fragments encoded by part or all of the sequence of Seq. ID No. 1 and which have morphogenic activity as defined herein. Specifically, as defined herein, a morphogen is a dimeric protein comprising a pair of polypeptide chains, wherein each polypeptide chain comprises at least the C-terminal six cysteine skeleton defined by residues 303 to 399 of Seq. ID No. 1 (or residues 335-431 of OP1, Seq. ID no. 3), including functionally equivalent arrangements of these cysteines (e.g., amino acid insertions or deletions which alter the linear arrangement of the cysteines in the sequence but not their relationship in the folded structure), such that, when the polypeptide chains are folded, the dimeric protein species comprising the pair of polypeptide chains has the appropriate three-dimensional structure, including the appropriate intra- or inter-chain disulfide bonds such that the protein is capable of acting as a morphogen as defined herein. Specifically, the morphogens generally are capable of all of the following biological functions in a morphogenically permissive environment: stimulating proliferation of progenitor cells; stimulating the differentiation of progenitor cells; stimulating the proliferation of differentiated cells; and supporting the growth and maintenance of differentiated cells.

In one aspect, the morphogens of this invention comprise a morphogenically active dimeric species comprising a pair of polypeptide chains, wherein at

least one of the polypeptide chains comprises the amino acid sequence defined by residues 303 to 399 of Seq. ID No. 1 including allelic, species and other amino acid sequence variants thereof. In preferred morphogens, at least one polypeptide chain comprises the sequence defined by residues 298-399, residues 261-399 or residues 264-399 of Seq. ID No. 1. Alternatively, the amino acid sequence of both polypeptide chains may be defined by part or all of the amino acid sequence of Seq. ID No. 1, including allelic, species and other amino acid sequence variants thereof, including naturally-occurring sequence or biosynthetically constructed variants, and chimeric constructs as described below. Where only one polypeptide chain is defined by the amino acid sequence of part or all of Seq. ID. No. 1, the other polypeptide chain preferably comprises at least the sequence defining the C-terminal six cysteine skeleton of any of the other known morphogen family members, including OP-1, OP-2, CBMP2A, CBMP2B, BMP3, BMP5, BMP6, Vgr-1, Vgl, 60A, DPP and GDF-1, described, for example, in US92/07432 (WO93/05751), including allelic, species and other amino acid sequence variants thereof, including chimeric variants. Other useful sequences include biosynthetic constructs, such as are described in U.S. Pat. No. 5,011,691.

In still another aspect of the invention, generic sequences are provided which accommodate the sequence identity of useful morphogens and incorporate OP-3's novel features.

In another aspect of the invention, morphogens of this invention comprise morphogenically active proteins encoded by part or all of the genetic sequence listed in Seq. ID No. 1, including allelic, species and other amino acid sequence variants thereof. In still another aspect, the invention comprises morphogens encoded by nucleic acids that hybridize to part or all of the pro region of the OP-3 protein, bases 120 to 848 of Seq ID No. 1, under stringent hybridization conditions. As used herein, "stringent hybridization conditions" are defined as hybridization in 40% formamide, 5 x SSPE, 5 x Denhardt's Solution, and 0.1% SDS at 37°C overnight, and washing in 0.1 x SSPE, 0.1% SDS at 50°C.

In one aspect of the invention, morphogenically active fragments of OP-3 are useful in the replacement of diseased or damaged tissue in a mammal, including, but not limited to, damaged lung tissue resulting from emphysema; cirrhotic tissue, including cirrhotic kidney or liver tissue; damaged heart or blood vessel tissue, as may result from cardiomyopathies and/or atherothrombotic or cardioembolic strokes; damaged stomach and other mucosal tissues of the gastrointestinal tract resulting from ulceric perforations and/or their repair; damaged nerve tissue as may result from physical injury, degenerative diseases such as Alzheimer's disease, multiple sclerosis, or strokes; damaged cartilage and bone tissue as may result from metabolic bone diseases and

other bone remodeling disorders; damaged dentin, periodontal and/or cementum tissue as may result from disease or mechanical injury; and in the replacement of damaged tissue as a result of inflammation and/or chronic inflammatory disease.

As provided herein, morphogenically active fragments of OP-3 are provided to a tissue-specific locus in vivo, to induce the developmental cascade of tissue morphogenesis at that site. Cells stimulated ex vivo by contact with OP-3 also may be provided to the tissue locus. In these cases the existing tissue provides the necessary matrix requirements, providing a suitable substratum or scaffold for the proliferating and differentiating cells in a morphogenically permissive environment, as well as providing the necessary signals for directing the tissue-specificity of the developing tissue. The proteins or stimulated cells also may be combined with a formulated matrix and implanted as a device at a locus in vivo. The formulated matrix should be a biocompatible, preferably biodegradable acellular matrix having the characteristics described below. Where the necessary signals for directing the tissue-specificity of the developing tissue are not provided endogenously, the matrix preferably also is tissue-specific.

In another aspect, the members of the morphogen protein family also can control the body's cellular and humoral inflammatory response to a foreign object or an initial tissue injury. In many instances, the loss of

tissue function results from the tissue destructive effects and the subsequent formation of scar tissue associated with the body's immune/inflammatory response to an initial or repeated injury to the tissue. The degree of scar tissue formation generally depends on the regenerative properties of the injured tissue, and on the degree and type of tissue damage. Thus, in another aspect, morphogenically active fragments of OP-3 may be used to prevent or to substantially inhibit the formation of scar tissue, including alleviating immune response-mediated tissue damage, by providing OP-3 or cells stimulated by exposure to OP-3 protein, to a newly injured tissue locus. The OP-3 protein also may be provided as a prophylactic, provided to a site in anticipation of tissue injury, such as part of a surgical or other clinical procedure likely to produce tissue damage, and to induce an inflammatory/immune response. In a particularly useful embodiment, OP-3 may be used as part of a transplant procedure, to enhance the tissue viability of the organ and/or tissue to be transplanted. The morphogen may be provided to the organ and/or tissue to be transplanted prior to harvest, during its transport, and/or during transplantation into the recipient host as described below.

OP-3 also may be used to increase or regenerate a mesenchymal progenitor or stem cell population <u>in vitro</u> or in a mammal. For example, progenitor cells may be isolated from an individual's bone marrow, stimulated

ex vivo with morphogenic OP-3 for a time and at a concentration sufficient to induce the cells to proliferate, and returned to the bone marrow. Other sources of progenitor cells that may be suitable include biocompatible cells obtained from a cultured cell line, stimulated in culture, and subsequently provided to the body. Alternatively, OP-3 may be provided by systemic (e.g., oral or parenteral) administration, or it may be injected or otherwise provided to a progenitor cell population in an individual to induce its mitogenic activity in vivo. For example, a morphogenically active fragment of OP-3 may be provided to the cells in vivo, e.g., by systemic injection, to induce mitogenic activity. Similarly, a particular population of hemopoietic stem cells may be increased by exposure to OP-3, for example by perfusing (plasmaphoresing) an individual's blood to extract the cells of interest, stimulating these cells ex vivo, and returning the stimulated cells to the blood.

It is anticipated that the ability to augment an individual's progenitor cell population will enhance existing methods for treating disorders resulting from a loss or reduction of a renewable cell population significantly. Two particularly significant applications include the treatment of blood disorders and diseases involving impaired or lost immune function.

The morphogens of this invention also can inhibit proliferation of epithelial cell populations. The ability to inhibit epithelial cell proliferation may be exploited to reduce tissue damage associated with psoriasis and dermatitis, and other inflammatory skin diseases, as well as ulcerative diseases of the gastrointestinal tract, such as, for example, in the healing of ulcers, including gastric ulcers, and the ulcerations induced in oral mucocitis and inflammatory bowel disease. Morphogens may be used to particular advantage as a cytoprotective agent in clinical therapies likely to effect proliferating epithelial populations, such as cancer radiotherapies and chemotherapies that typically induce oral mucositis, hair loss and/or skin disorders.

In another aspect of the invention, morphogenic OP-3 may be used to support the growth and maintenance of differentiated cells, inducing existing differentiated cells to continue expressing their phenotype. It is anticipated that this activity will be particularly useful in the treatment of tissue disorders where loss of function is caused by reduced or lost metabolic function in which cells become senescent or quiescent, such as may occur in aging cells and/or may be manifested in osteoporosis and a number of nerve degenerative diseases, including Alzheimer's disease. Application of OP-3 directly to the cells to be treated, or providing it systemically, as by oral or parenteral administration, can stimulate

these cells to continue expressing their phenotype, thereby significantly reversing the effects of the dysfunction. In addition, a morphogenically active fragment of OP-3 also may be used in gene therapy protocols to stimulate the growth of quiescent cells, thereby potentially enhancing the ability of these cells to incorporate exogenous DNA.

In yet another aspect of the invention, a morphogenically active fragment of OP-3 also may be used to induce "redifferentiation" of cells that have strayed from their differentiation pathway, such as can occur during tumorgenesis. It is anticipated that this activity will be particularly useful in treatments to reduce or substantially inhibit the growth of neoplasms. The method also is anticipated to induce the de- and/or re-differentiation of these cells. As described supra, a morphogenically active OP-3 fragment may be provided to the cells directly or systemically, stimulating these cells to revert back to a morphology and phenotype characteristic of untransformed cells.

In still another aspect of the invention, OP-3 may be used to stimulate cell adhesion molecule (CAM) expression levels in a cell. CAMs are molecules defined as carrying out cell-cell interactions necessary for tissue formation. CAMs are believed to play a fundamental regulatory role in tissue development, including tissue boundary formation,

embryonic induction and migration, and tissue stabilization and regeneration. Altered CAM levels have been implicated in a number of tissue disorders, including congenital defects, neoplasias, and degenerative diseases.

In particular, N-CAM expression is associated with normal neuronal cell development and differentiation, including retinal formation, synaptogenesis, and nerve-muscle tissue adhesion. Inhibition of one or more of the N-CAM isoforms is known to prevent proper tissue development. Altered N-CAM expression levels also are associated with neoplasias, including neuroblastomas (see infra), as well as with a number of neuropathies, including normal pressure hydrocephalous and type II schizophrenia. Application of the morphogen directly to the cells to be treated, or providing the morphogen to the mammal systemically, for example, parenterally, or indirectly by oral administration, may be used to induce cellular expression of one or more CAMs, particularly N-CAMs and L1.

CAMs also have been postulated as part of a morphoregulatory pathway whose activity is induced by a to date unidentified molecule (See, for example, Edelman, G.M. (1986) Ann. Rev. Cell Biol., 2:81-116). Without being limited to any given theory, the morphogens described herein may act as inducers of this pathway.

The matrices utilized in the methods of the invention may be derived from organ-specific tissue, or they may be formulated synthetically. In one embodiment of the invention, when OP-3 (or a collection of progenitor cells stimulated by OP-3) is provided at a tissue-specific locus, e.g., by systemic administration, implantation or injection at a tissue-specific locus, the existing tissue at that locus, whether diseased or damaged, has the capacity of acting as a suitable matrix or scaffold for the differentiation and proliferation of migrating progenitor cells. Alternatively, a formulated matrix may be provided externally together with the stimulated progenitor cells or morphogenically active OP-3 fragment, as may be necessary when the extent of injury sustained by the damaged tissue is large. The matrix should be a biocompatible, suitably modified acellular matrix having dimensions such that it allows the differentiation and proliferation of migratory progenitor cells, and is capable of providing a morphogenically permissive environment. The matrix also preferably allows cellular attachment and is biodegradable. Where the necessary tissue-directing signals can not be provided endogenously, the matrix preferably also is tissue-specific.

Formulated matrices may be generated from dehydrated organ-specific tissue prepared, for example, by treating the tissue with solvents to substantially remove the intracellular, non-structural components

from the tissue. Alternatively, the matrix may be formulated synthetically using a biocompatible, preferably in vivo biodegradable, structural molecule, and may be formulated with suitable tissue-specific cell attachment factors. The molecule may be a naturally occurring one such as collagen, laminin or hyaluronic acid, or a synthetic polymer comprising, for example, polylactic acid, polybutyric acid, polyglycolic acid, and copolymers thereof. Currently preferred structural polymers comprise tissue-specific collagens. Currently preferred cell attachment factors include glycosaminoglycans and proteoglycans. The matrix further may be treated with an agent or agents to increase the number of pores and micropits on its surfaces, so as to enhance the influx, proliferation and differentiation of migratory progenitor cells from the body of the mammal.

The invention thus relates to compositions and methods for the use of morphogenically active fragments of OP-3, a novel species variant of the generic family of morphogens disclosed in USSN 667,274 and USSN 752,764, as a tissue morphogen. Morphogenically active OP-3 and protein fragments can be isolated from naturally-occurring sources, or they may be constructed biosynthetically using conventional recombinant DNA technology. Active OP-3 useful in the compositions and methods of this invention may include forms having varying glycosylation patterns, varying N-termini and active truncated forms, e.g., produced by recombinant

DNA techniques. Active OP-3 proteins also include chimeric constructs as described below, comprising both an OP-3 active domain and a non-OP-3 sequence as, for example, the pro domain and/or the N-terminal region of the mature protein. OP-3 protein can be expressed from intact or truncated cDNA or from synthetic DNAs in procaryotic or eucaryotic host cells, and purified, cleaved, refolded, and dimerized to form morphogenically active compositions. Useful host cells include procaryotes, including E. coli, and eucaryotic cells, including mammalian cells, such as CHO, COS, melanoma or BSC cells, or the insect/baculovirus system. Thus recombinant DNA techniques may be utilized to produce large quantities of OP-3 capable of inducing tissue-specific cell differentiation and tissue morphogenesis in a variety of mammals, including humans.

Brief Description of the Drawings

Figure 1 is a nucleotide sequence comparison of the mouse cDNA sequence of OP-2 and OP-3. Exon boundaries are indicated by bars beneath the sequence; diamonds indicate nucleotide differences within exons 2 and 3; and

Figure 2 is an immunoblot comparing mammalian cell expression of an OP1/OP3 chimeric protein construct (lanes 4-8) with that of authentic, recombinant OP1 (lane 1).

Detailed Description

The invention provides a novel genetic sequence, mOP-3, encoding a novel protein, OP-3, having morphogenic properties. The genetic sequence originally was identified in a mouse cDNA library, and the invention provides methods for identifying and isolating the gene from other species. As will be appreciated by those skilled in the art, the methods described herein also may be used to isolate the OP-3 gene from other libraries, including genomic libraries. The invention also provides means for producing the OP-3 genetic sequence and the encoded protein. The invention further provides methods and compositions for inducing the developmental cascade of tissue morphogenesis in a mammal utilizing morphogenically active fragments of OP-3. The methods and compositions provided herein may be utilized in a range of applications, including stimulating the proliferation and/or differentiation of progenitor cells and inducing the repair and regeneration of damaged tissue. The morphogenic OP-3 species of the invention are novel members of the family of morphogens disclosed in US92/01968 (WO92/15323) and US92/07432 (WO93/05751). As described herein, OP-3 may be isolated from natural sources or constructed biosynthetically utilizing conventional recombinant DNA technology or constructed synthetically using standard chemical techniques.

Morphogenically active fragments of OP-3 are useful for initiating and maintaining the tissue-specific developmental cascade in a variety of tissues,

including, but not limited to, bone, cartilage, dentin, neural tissue, liver, periodontal ligament, cementum, lung, heart, kidney and numerous tissues of the gastrointestinal tract. When combined with naive mesenchymal progenitor cells as disclosed herein, OP-3 can induce the proliferation and differentiation of these progenitor cells. In the presence of appropriate tissue-specific signals to direct the differentiation of these cells, and a morphogenically permissive environment, OP-3 is capable of reproducing the cascade of cellular and molecular events that occur during embryonic development to yield functional tissue. For example, the protein can induce the de novo formation of cartilage and endochondral bone, including inducing the proliferation and differentiation of progenitor cells into chondrocytes and osteoblasts, inducing appropriate mineralization and bone remodeling, inducing formation of an appropriate bone tissue vascular supply and inducing formation of differentiated bone marrow (see Example 7 below.)

Provided below is a detailed description of the nucleic acid and amino acid sequences which describe OP-3 proteins useful in the compositions and methods of this invention, including a description of how to make them, and methods and means for their therapeutic administration. Also provided are numerous, nonlimiting examples which (1) illustrate the suitability of these proteins as tissue morphogens and therapeutic agents, and (2) provide assays with which to test the morphogens encompassed by the invention in

different tissues. Also provided in Example 9 is a method for screening compounds to identify morphogen stimulating agents capable of stimulating endogenous OP-3 expression and/or secretion. OP-3 stimulating agents then may be used in any of the therapeutic applications described herein in place of, or in addition to, OP-3 protein administration.

I. Useful Morphogens

As defined herein a protein is morphogenic if it is capable of inducing the developmental cascade of cellular and molecular events that culminate in the formation of new, organ-specific tissue and comprises at least the conserved C-terminal six cysteine skeleton or its functional equivalent (see supra). Specifically, the morphogens generally are capable of all of the following biological functions in a morphogenically permissive environment: stimulating proliferation of progenitor cells; stimulating the differentiation of progenitor cells; stimulating the proliferation of differentiated cells; and supporting the growth and maintenance of differentiated cells. Details of how the morphogen family of proteins described herein first were identified, as well as a description of how to make, use and test them for morphogenic activity are disclosed, for example, in international application US92/01968 (WO92/15323). As disclosed therein, the morphogens may be purified from naturally-sourced material or recombinantly produced

from procaryotic or eucaryotic host cells, preferably as described therein. Alternatively, novel morphogenic sequences may be identified following the procedures disclosed therein.

Particularly useful morphogens identified to date include OP-1, OP-2, CBMP2A and CBMP2B (the morphogenically active domains of proteins referred to in the art as BMP2A and BMP2B, or BMP2 and BMP4, respectively), BMP3, BMP5, BMP6, Vgr-1, GDF-1, Vgl, DPP and 60A, including their allelic and species variants, as well as other amino acid sequence variants, including chimeric morphogens. Morphogenically active biosynthetic constructs such as those disclosed in U.S. Pat. No. 5,011,691, (e.g., COP-1, COP-3, COP-4, COP-5, COP-7, and COP-16) also are envisioned to be useful.

The novel morphogen OP-3 and its genetic sequence now have been identified. The OP-3 proteins useful in the invention include any morphogenically active fragment of the OP-3 amino acid sequence present in Seq. ID No. 1, or allelic, species or other amino acid sequence variants thereof. The morphogenically active fragment of OP-3 also may include any morphogenically active protein encoded by part or all of the nucleic acid sequence presented in Seq. ID No. 1. The morphogenic protein also may comprise a protein encoded by part or all of a nucleic acid which hybridizes to at least part of the nucleic acid sequence encoding the "pro" region of the OP-3 protein, e.g., bases 120-848 of Seq. ID No. 1, under stringent conditions.

The mOP-3 gene encodes a protein ("mOP-3") first expressed as an immature translation product that is 399 amino acids in length. This precursor form, referred to herein as the "prepro" form, (Seq. ID. No. 1, amino acid residues 1-399) includes an N-terminal signal peptide sequence, typically less than about 20 residues, followed by a "pro" domain that is cleaved to yield the mature sequence. The "pro" form of the protein includes the pro domain and the mature domain, and forms a soluble species that appears to be the primary form secreted from cultured mammalian cells. The signal peptide, anticipated to include residues 1-17 for mOP3, is cleaved rapidly upon translation, at a cleavage site that can be predicted in a given sequence using the method of Von Heijne ((1986) Nucleic Acids Research 14:4683-4691). The preferred form of morphogenically active OP-3 protein comprises a processed sequence, including fragments thereof, appropriately dimerized and disulfide bonded. Where a soluble form of the protein is desired, the protein preferably comprises both the mature domain, or an active portion thereof, and part or all of the pro domain.

By amino acid sequence homology with other, known morphogens, the pro domain likely is cleaved at residues 257-260 of Seq. ID No. 1, which represent the canonical Arg-Xaa-Xaa-Arg cleavage site, to yield a mature sequence 139 amino acids in length (Seq. ID No. 1, residues 261-399). Alternatively, the pro domain may be cleaved at residues 260-263 to yield a

shorter sequence 135 amino acids in length (Seq. ID No. 1, amino acid residues 264-399). All morphogens, including OP-1, OP-2 and the OP-3 proteins disclosed herein, comprise at least a conserved six cysteine skeleton in the amino acid sequence C-terminal domain and, preferably, a conserved seven cysteine skeleton (see, for example, US92/01968 (WO92/15323). The conserved six cysteine skeleton in mOP-3 (Seq. ID No. 1) is defined by amino acid residues 303-399; the conserved seven cysteine skeleton is defined by amino acid residues 298-399. In addition to the conserved six cysteine skeleton found in known morphogen family members including OP-1, OP-2, CBMP2A, CBMP2B, BMP3, BMP5, BMP6, Vgr-1, Vgl, 60A, DPP and GDF-1, described, for example, in PCT/US92/07432 (WO93/05751), the OP-3 proteins, like the OP-2 proteins, also has one additional cysteine residue (residue 338 of Seq. ID No. 1) in the conserved C-terminal domain.

The mature sequence of OP-3 shares significant amino acid sequence homology with the morphogens identified to date. Specifically, the seven cysteine fragment shows greater than 79% amino acid identity with the corresponding mOP-2 and hOP-2 sequences, and greater than 66% identity with the corresponding OP-1 sequences. Like OP-2, OP-3 has an eighth cysteine within the seven cysteine domain (e.g., at position 338 of Seq. ID No. 1). In addition, OP-3 is unique among

the morphogens identified to date in that the residue at position 9 in the conserved seven cysteine domain (e.g., residue 315 of Seq. ID No. 1) is a serine, whereas other morphogens typically have a tryptophan at this location (see Table I below, and Table II in PCT/US92/07358 (WO93/04692), for example.)

As used herein, "amino acid sequence homology" is understood to mean amino acid sequence similarity, and homologous sequences share identical or similar amino acids, where similar amino acids are conserved amino acids as defined by Dayoff et al., Atlas of Protein Sequence and Structure; vol.5, Suppl.3, pp.345-362 (M.O. Dayoff, ed., Nat'l BioMed. Research Fdn., Washington D.C. 1978.) Thus, a candidate sequence sharing 70% amino acid homology with a reference sequence requires that, following alignment of the candidate sequence with the reference sequence, 70% of the amino acids in the candidate sequence are identical to the corresponding amino acid in the reference sequence, or constitute a conserved amino acid change thereto. "Amino acid sequence identity" is understood to require identical amino acids between two aligned sequences. Thus, a candidate sequence sharing 60% amino acid identity with a reference sequence requires that, following alignment of the candidate sequence with the reference sequence, 60% of the amino acids in the candidate sequence are identical to the corresponding amino acid in the reference sequence.

As used herein, all homologies and identities calculated use OP-3 as the reference sequence. Also as used herein, sequences are aligned for homology and identity calculations using the method of Needleman et al. (1970) J.Mol. Biol. 48:443-453 and identities calculated by the Align program (DNAstar, Inc.) In all cases, internal gaps and amino acid insertions in the candidate sequence as aligned are ignored when making the homology/identity calculation.

Thus, useful OP-3 variants include, but are not limited to, amino acid sequences derived from Seq. ID No. 1 and wherein the cysteine at position 338 is replaced with another amino acid, preferably a tyrosine, histidine, isoleucine or serine and conservative substitutions thereof, e.g., such as defined by Dayoff et al., Atlas of Protein Sequence and Structure; vol. 5, Suppl. 3, pp.345-362 (M.O. Dayoff, ed., Nat'l BioMed. Research Fdn., Washington D.C. 1979.). Still other useful OP-3 variants include proteins wherein the serine at position 315 is replaced with another amino acid, preferably a tryptophan and conservative substitutions thereof.

Generic Sequence 7 (Seq. ID No. 12) and Generic Sequence 8 (Seq. ID No. 13) disclosed below, accommodate the homologies shared among perferred morphogen protein family members identified to date, including OP-1, OP-2, OP-3, CBMP2A, CBMP2B, BMP3, 60A, DPP, Vg1, BMP5, BMP6, Vrg-1, and GDF-1. The amino acid sequences for these proteins are described herein (see

Sequence Listing and Table I below) and/or in the art, as well as in PCT publication US 92/07358, filed August 28, 1992, for example. The generic sequences include both the amino acid identity shared by these sequences in the C-terminal domain, defined by the six and seven cysteine skeletons (Generic Sequences 7 and 8, respectively), as well as alternative residues for the variable positions within the sequence. The generic sequences allow for an additional cysteine at position 41 (Generic Sequence 7) or position 46 (Generic Sequence 8), providing an appropriate cysteine skeleton where inter- or intramolecular disulfide bonds can form, and containing certain critical amino acids which influence the tertiary structure of the proteins.

Generic Sequence 7

Leu Xaa Xaa Xaa Phe
1 5

Xaa Xaa Xaa Gly Trp Xaa Xaa Xaa Xaa
10

Xaa Xaa Pro Xaa Xaa Xaa Xaa Ala
15 20

Xaa Tyr Cys Xaa Gly Xaa Cys Xaa
25 30

Xaa Pro Xaa Xaa Xaa Xaa Xaa
35
Xaa Xaa Xaa Asn His Ala Xaa Xaa
40 45
Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa
50
Xaa Xaa Xaa Xaa Xaa Xaa Xaa Cys
55 60
Cys Xaa Pro Xaa Xaa Xaa Xaa Xaa
65
Xaa Xaa Xaa Leu Xaa Xaa Xaa
70 75
Xaa Xaa Xaa Xaa Val Xaa Leu Xaa
80
Xaa Xaa Xaa Xaa Met Xaa Val Xaa
85 90
Xaa Cys Xaa Cys Xaa
95

wherein each Xaa is independently selected from a group of one or more specified amino acids defined as follows: "Res." means "residue" and Xaa at res.2 = (Tyr or Lys); Xaa at res.3 = Val or Ile); Xaa at res.4 = (Ser, Asp or Glu); Xaa at res.6 = (Arg, Gln, Ser, Lys or Ala); Xaa at res.7 = (Asp or Glu); Xaa at res.8 =

(Leu, Val or Ile); Xaa at res.11 = (Gln, Leu, Asp, His, Asn or Ser); Xaa at res.12 = (Asp, Arg, Asn or Glu); Xaa at res. 13 = (Trp or Ser); Xaa at res.14 = (Ile or Val); Xaa at res.15 = (Ile or Val); Xaa at res.16 (Ala or Ser); Xaa at res.18 = (Glu, Gln, Leu, Lys, Pro or Arg); Xaa at res.19 = (Gly or Ser); Xaa at res.20 = (Tyr or Phe); Xaa at res.21 = (Ala, Ser, Asp, Met, His, Gln, Leu or Gly); Xaa at res.23 = (Tyr, Asn or Phe); Xaa at res.26 = (Glu, His, Tyr, Asp, Gln, Ala or Ser); Xaa at res.28 = (Glu, Lys, Asp, Gln or Ala); Xaa at res.30 = (Ala, Ser, Pro, Gln, Ile or Asn); Xaa at res.31 = (Phe, Leu or Tyr); Xaa at res.33 = (Leu, Val or Met); Xaa at res.34 = (Asn, Asp, Ala, Thr or Pro); Xaa at res.35 = (Ser, Asp, Glu, Leu, Ala or Lys); Xaa at res.36 = (Tyr, Cys, His, Ser or Ile); Xaa at res.37 = (Met, Phe, Gly or Leu); Xaa at res.38 = (Asn, Ser or Lys); Xaa at res.39 = (Ala, Ser, Gly or Pro); Xaa at res.40 = (Thr, Leu or Ser); Xaa at res.44 = (Ile, Val or Thr); Xaa at res.45 = (Val, Leu, Met or Ile); Xaa at res.46 = (Gln or Arg); Xaa at res.47 = (Thr, Ala or Ser); Xaa at res.48 = (Leu or Ile); Xaa at res.49 = (Val or Met); Xaa at res.50 = (His, Asn or Arg); Xaa at res.51 = (Phe, Leu, Asn, Ser, Ala or Val); Xaa at res.52 = (Ile, Met, Asn, Ala, Val, Gly or Leu); Xaa at res.53 = (Asn, Lys, Ala, Glu, Gly or Phe); Xaa at res.54 = (Pro, Ser or Val); Xaa at res.55 = (Glu, Asp, Asn, Gly, Val, Pro or Lys); Xaa at res.56 = (Thr, Ala, Val, Lys, Asp, Tyr, Ser, Gly, Ile or His); Xaa at res.57 = (Val, Ala or Ile); Xaa at res.58 = (Pro or Asp); Xaa at res.59 = (Lys, Leu or Glu); Xaa at res.60 = (Pro, Val or Ala); Xaa at res.63 = (Ala or

Val); Xaa at res.65 = (Thr, Ala or Glu); Xaa at res.66 = (Gln, Lys, Arg or Glu); Xaa at res.67 = (Leu, Met or Val); Xaa at res.68 = (Asn, Ser, Asp or Gly); Xaa at res.69 = (Ala, Pro or Ser); Xaa at res.70 = (Ile, Thr, Val or Leu); Xaa at res.71 = (Ser, Ala or Pro); Xaa at res.72 = (Val, Leu, Met or Ile); Xaa at res.74 = (Tyr or Phe); Xaa at res.75 = (Phe, Tyr, Leu or His); Xaa at res.76 = (Asp, Asn or Leu); Xaa at res.77 = (Asp, Glu, Asn, Arg or Ser); Xaa at res.78 = (Ser, Gln, Asn, Tyr or Asp); Xaa at res.79 = (Ser, Asn, Asp, Glu or Lys); Xaa at res.80 = (Asn, Thr or Lys); Xaa at res.82 = (Ile, Val or Asn); Xaa at res.84 = (Lys or Arg); Xaa at res.85 = (Lys, Asn, Gln, His, Arg or Val); Xaa at res.86 = (Tyr, Glu or His); Xaa at res.87 = (Arg, Gln, Glu or Pro); Xaa at res.88 = (Asn, Glu, Trp or Asp); Xaa at res.90 = (Val, Thr, Ala or Ile); Xaa at res.92 = (Arg, Lys, Val, Asp, Gln or Glu); Xaa at res.93 = (Ala, Gly, Glu or Ser); Xaa at res.95 = (Gly or Ala) and Xaa at res.97 = (His or Arg).

As described above, Generic Sequence 8 (Seq. ID No. 13) includes all of Generic Sequence 7 and in addition includes the following sequence at its N-terminus:

Cys Xaa Xaa Xaa Xaa
1 5

Accordingly, beginning with residue 7, each "Xaa" in Generic Seq. 8 is a specified amino acid defined as for Generic Seq. 7, with the distinction that each residue number described for Generic Sequence 7 is

shifted by five in Generic Seq. 8. Thus, "Xaa at res.2 =(Tyr or Lys)" in Gen. Seq. 7 refers to Xaa at res. 7 in Generic Seq. 8. In Generic Seq. 8, Xaa at res.2 = (Lys, Arg, Ala or Gln); Xaa at res.3 = (Lys, Arg or Met); Xaa at res.4 = (His, Arg or Gln); and Xaa at res.5 = (Glu, Ser, His, Gly, Arg, Pro, Thr, or Tyr).

Table I, set forth below, compares the C-terminal amino acid sequences defining the seven cysteine skeleton of human OP-1, mouse OP-1, human OP-2, mouse OP-2, and mouse OP-3 (mOP-3, Seq. ID No. 1). In the table, the sequences are aligned essentially following the method of Needleman et al. (1970) J. Mol. Biol., 48:443-453, calculated using the Align Program (DNAstar, Inc.) In the table, three dots indicates that the amino acid in that position is the same as the amino acid in hOP-1. Three dashes indicate that no amino acid is present in that position, and are included for purposes of illustrating homologies. As is apparent from the following amino acid sequence comparisons, significant amino acid sequence homology exists between mouse OP-3 and mouse and human OP-1 and OP-2.

TABLE I

hOP-1	Cys	Lys	Lys	His	Glu	Leu	Tyr	Val
mOP-1	...	...	...	...	...	...	...	...
hOP-2	...	Arg	Arg	...	...	...	...	...
mOP-2	...	Arg	Arg	...	...	...	...	...
mOP-3	...	Arg	Arg	...	...	...	...	...
	1				5			

	hOP-1	Ser	Phe	Arg	Asp	Leu	Gly	Trp	Gln	Asp
	mOP-1	...	...	...	...	...	...	...	...	...
	hOP-2	...	...	Gln	...	...	...	...	Leu	...
	mOP-2	...	...	...	...	...	...	...	Leu	...
5	mOP-3	...	...	...	...	...	...	...	Leu	...

10 15

	hOP-1	Trp	Ile	Ile	Ala	Pro	Glu	Gly	Tyr	Ala
10	mOP-1	...	...	...	...	...	...	...	...	...
	hOP-2	...	Val	...	...	...	Gln	...	...	Ser
	mOP-2	...	Val	...	...	...	Gln	...	...	Ser
	mOP-3	Ser	Val	...	...	...	Gln	...	...	Ser

15 20 25

	hOP-1	Ala	Tyr	Tyr	Cys	Glu	Gly	Glu	Cys	Ala
	mOP-1	...	...	...	...	...	...	...	...	...
20	hOP-2	...	...	...	...	...	...	...	...	Ser
	mOP-2	...	...	...	...	...	...	...	...	...
	mOP-3	...	...	...	...	Ala	...	...	...	Ile

30 35

25

	hOP-1	Phe	Pro	Leu	Asn	Ser	Tyr	Met	Asn	Ala
	mOP-1	...	...	...	...	...	...	...	...	...
	hOP-2	...	...	...	Asp	...	Cys	...	...	...
	mOP-2	...	...	...	Asp	...	Cys	...	...	...
30	mOP-3	Tyr	...	...	...	...	Cys	...	...	Ser

40

	hOP-1	Thr	Asn	His	Ala	Ile	Val	Gln	Thr	Leu
35	mOP-1	...	...	...	...	...	...	...	...	...
	hOP-2	...	...	...	...	...	Leu	...	Ser	...
	mOP-2	...	...	...	...	...	Leu	...	Ser	...
	mOP-3	...	...	...	...	Thr	Met	...	Ala	...

40 45 50

	hOP-1	Val	His	Phe	Ile	Asn	Pro	Glu	Thr	Val
	mOP-1	...	...	...	...	...	...	Asp	...	...
45	hOP-2	...	...	Leu	Met	Lys	...	Asn	Ala	...
	mOP-2	...	...	Leu	Met	Lys	...	Asp	Val	...
	mOP-3	...	...	Leu	Met	Lys	...	Asp	Ile	Ile

55 60

50

hOP-1	Pro	Lys	Pro	Cys	Cys	Ala	Pro	Thr	Gln
mOP-1	...	...	...	...	...	...	...	...	...
hOP-2	...	...	Ala	...	...	...	...	...	Lys
mOP-2	...	...	Ala	...	...	...	...	...	Lys
mOP-3	...	...	Val	...	...	Val	...	...	Glu
			65					70	
hOP-1	Leu	Asn	Ala	Ile	Ser	Val	Leu	Tyr	Phe
mOP-1	...	...	...	...	...	...	...	...	...
hOP-2	...	Ser	...	Thr	...	...	...	...	Tyr
mOP-2	...	Ser	...	Thr	...	...	...	...	Tyr
mOP-3	...	Ser	...	...	...	Leu	...	...	Tyr
				75					80
hOP-1	Asp	Asp	Ser	Ser	Asn	Val	Ile	Leu	Lys
mOP-1	...	...	...	...	...	...	...	...	...
hOP-2	...	Ser	...	Asn	...	...	...	...	Arg
mOP-2	...	Ser	...	Asn	...	...	...	...	Arg
mOP-3	...	Arg	Asn	Asn	...	...	...	...	Arg
					85				
hOP-1	Lys	Tyr	Arg	Asn	Met	Val	Val	Arg	
mOP-1	...	...	...	...	...	...	...	...	
hOP-2	...	His	...	...	...	...	...	Lys	
mOP-2	...	His	...	...	...	...	...	Lys	
mOP-3	Arg	Glu	...	...	...	...	...	Gln	
	90					95			
hOP-1	Ala	Cys	Gly	Cys	His				
mOP-1	...	...	...	...	...				
hOP-2	...	...	...	...	...				
mOP-2	...	...	...	...	...				
mOP-3	...	...	...	...	...				
			100						

II. Formulations and Methods for Administering OP-3 Protein as Therapeutic Agents

II.A OP-3 Protein Considerations

The morphogens described herein may be provided to an individual by any suitable means, preferably directly or systemically, e.g., parenterally or orally. Where the morphogen is to be provided directly (e.g., locally, as by injection, to a desired tissue site), or parenterally, such as by intravenous, subcutaneous, intramuscular, intraorbital, ophthalmic, intraventricular, intracranial, intracapsular, intraspinal, intracisternal, intraperitoneal, buccal, rectal, vaginal, intranasal or by aerosol administration, the morphogen preferably comprises part of an aqueous solution. The solution is physiologically acceptable so that in addition to delivery of the desired morphogen to the patient, the solution does not otherwise adversely affect the patient's electrolyte and volume balance. The aqueous medium for the morphogen thus may comprise normal physiologic saline (0.9% NaCl, 0.15M), pH 7-7.4. The aqueous solution containing the morphogen can be made, for example, by dissolving the protein in 50% ethanol, or acetonitrile containing 0.1% trifluoroacetic acid (TFA) or 0.1% HCl, or equivalent solvents. One volume of the resultant solution then is added, for example, to ten volumes of phosphate buffered saline (PBS), which further may include 0.1-0.2% human serum albumin (HSA). The resultant solution preferably is vortexed extensively.

If desired, a given morphogen may be made more soluble by association with a suitable molecule. For example, association of the mature dimer with the pro domain of the morphogen increases solubility of the protein significantly. For example,the pro form of OP-3 comprises a species that is soluble in physiologically buffered solutions. In fact, the endogenous protein is thought to be transported (e.g., secreted and circulated) to particular tissues in this form. This soluble form of the protein may be obtained from the culture medium of morphogen-secreting mammalian cells. Alternatively, a soluble species may be formulated by complexing the mature dimer (or an active fragment thereof) with part or all of a pro domain. Another molecule capable of enhancing solubility and particularly useful for oral administrations, is casein. For example, addition of 0.2% casein increases solubility of the mature active form of OP-1 by 80%. Other components found in milk and/or various serum proteins also may be useful.

Useful solutions for oral or parenteral administration may be prepared by any of the methods well known in the pharmaceutical art, described, for example, in Remington's Pharmaceutical Sciences, (Gennaro, A., ed.), Mack Pub., 1990. Formulations may include, for example, polyalkylene glycols such as polyethylene glycol, oils of vegetable origin, hydrogenated naphthalenes, and the like. Formulations for direct administration, in particular, may include glycerol and other compositions of high viscosity.

Biocompatible, preferably bioresorbable polymers, including, for example, hyaluronic acid, collagen, tricalcium phosphate, polybutyrate, polylactide, polyglycolide and lactide/glycolide copolymers, may be useful excipients to control the release of the morphogen in vivo.

Other potentially useful parenteral delivery systems for these morphogens include ethylene-vinyl acetate copolymer particles, osmotic pumps, implantable infusion systems, and liposomes. Formulations for inhalation administration may contain as excipients, for example, lactose, or may be aqueous solutions containing, for example, polyoxyethylene-9-lauryl ether, glycocholate and deoxycholate, or oily solutions for administration in the form of nasal drops, or as a gel to be applied intranasally.

Alternatively, the morphogens described herein may be administered orally. Oral administration of proteins as therapeutics generally is not practiced as most proteins readily are degraded by digestive enzymes and acids in the mammalian digestive system before they can be absorbed into the bloodstream. However, the morphogens described herein typically are acid-stable and protease-resistant (see, for example, U.S. Pat. No. 4,968,590.) In addition, at least one morphogen, OP-1, has been identified in bovine mammary gland extract, colostrum and milk, as well as saliva. Moreover, the OP-1 purified from mammary gland extract is morphogenically active. For example, this protein

induces endochondral bone formation in mammals when implanted subcutaneously in association with a suitable matrix material, using a standard in vivo bone assay, such as is disclosed in U.S. Pat. No. 4,968,590. In addition, endogenous morphogen also is detected in human serum. These findings indicate that oral and parenteral administration are viable means for administering morphogens to an individual. Moreover, while the mature forms of certain morphogens described herein typically are sparingly soluble, the morphogen form found in milk (and mammary gland extract and colostrum) is readily soluble, probably by association of the mature, morphogenically active form with the pro domain of the intact sequence and/or by association with one or more milk components. Accordingly, the compounds provided herein also may be associated with molecules capable of enhancing their solubility in vitro or in vivo, including, for example, part or all of a morphogen pro domain, as described below, and casein, as described above.

The compounds provided herein also may be associated with molecules capable of targeting the morphogen to a desired tissue. For example, tetracycline and diphosphonates (bisphosphonates) are known to bind to bone mineral, particularly at zones of bone remodeling, when they are provided systemically in a mammal. Accordingly, these molecules may be included as useful agents for targeting OP-3 to bone tissue. Alternatively, an antibody or other binding protein that interacts specifically with a surface molecule on

the desired target tissue cells also may be used. Such targeting molecules further may be covalently associated to the morphogen, e.g., by chemical crosslinking, or by using standard genetic engineering means to create, for example, an acid labile bond such as an Asp-Pro linkage. Useful targeting molecules may be designed, for example, using the single chain binding site technology disclosed, for example, in U.S. Pat. No. 5,091,513.

As described above, the morphogen family members share significant sequence homology in the C-terminal active domains. By contrast, the sequences diverge significantly in the sequences which define the pro domain and the N-terminal 39 amino acids of the mature protein. Accordingly, the pro domain and/or N-terminal sequence may be morphogen-specific. As described above, it also is known that the various morphogens identified to date are differentially expressed in the different tissues. Accordingly, without being limited to any given theory, it is likely that, under natural conditions in the body, selected morphogens typically act on a given tissue. Accordingly, part or all of morphogen-specific sequences may serve as targeting molecules for the morphogens described herein. For example, the pro domains may interact specifically with one or more molecules at the target tissue to direct the morphogen associated with the pro domain to that tissue. Thus, another useful targeting molecule for targeting OP-3 to bone tissue, for example, may include part or all of a morphogen-specific sequence, such as

part or all of a pro domain and/or the N-terminus of the mature protein. Particularly useful are the morphogen-specific sequences of OP-1, BMP2 or BMP4, all of which proteins are found naturally associated with bone tissue (see, for example, US Pat. No. 5,011,691). Alternatively, the morphogen-specific sequences of GDF-1 may be used to target morphogenic OP-3 to nerve tissue, particularly brain tissue where GDF-1 appears to be primarily expressed (see, for example, Lee, (1991) PNAS, 88:4250-4254. As described above, pro forms of the proteins may be obtained from the culture medium of morphogen-secreting mammalian cells. Alternatively, a suitable species may be formulated by complexing the mature dimer (or an active fragment thereof) with part or all of a pro domain. Chimeric OP-3 proteins comprising, for example, non-OP-3 pro domains and/or non-OP-3 N-termini, may be synthesized using standard recombinant DNA methodology and/or automated chemical nucleic acid synthesis methodology well described in the art and as disclosed below.

Finally, the OP-3 proteins provided herein may be administered alone or in combination with other molecules known to have a beneficial effect on tissue morphogenesis, including molecules capable of tissue repair and regeneration and/or inhibiting inflammation. Examples of useful cofactors for stimulating bone tissue growth in osteoporotic individuals, for example, include but are not limited to, vitamin D_3, calcitonin, prostaglandins, parathyroid hormone, dexamethasone, estrogen and IGF-I or IGF-II. Useful cofactors for

nerve tissue repair and regeneration may include nerve growth factors. Other useful cofactors include symptom-alleviating cofactors, including antiseptics, antibiotics, antiviral and antifungal agents and analgesics and anesthetics.

The compounds provided herein can be formulated into pharmaceutical compositions by admixture with pharmaceutically acceptable nontoxic excipients and carriers. As noted above, such compositions may be prepared for parenteral administration, particularly in the form of liquid solutions or suspensions; for oral administration, particularly in the form of tablets or capsules; or intranasally, particularly in the form of powders, nasal drops or aerosols. Where adhesion to a tissue surface is desired the composition may include the morphogen dispersed in a fibrinogen-thrombin composition or other bioadhesive such as is disclosed, for example in PCT US91/09275, (WO92/10567). The composition then may be painted, sprayed or otherwise applied to the desired tissue surface.

The compositions can be formulated for parenteral or oral administration to humans or other mammals in therapeutically effective amounts, e.g., amounts which provide appropriate concentrations of OP-3 to target tissue for a time sufficient to induce morphogenesis, including particular steps thereof, as described above.

Where OP-3 is to be used as part of a transplant procedure, the morphogen may be provided to the living tissue or organ to be transplanted prior to removal of tissue or organ from the donor. OP-3 may be provided to the donor host directly, as by injection of a formulation comprising OP-3 into the tissue, or indirectly, e.g., by oral or parenteral administration, using any of the means described above.

Alternatively or, in addition, once removed from the donor, the organ or living tissue may be placed in a preservation solution containing OP-3. In addition, the recipient also preferably is provided with the morphogen just prior to, or concommitant with, transplantation. In all cases, OP-3 may be administered directly to the tissue at risk, as by injection to the tissue, or it may be provided systemically, either by oral or parenteral administration, using any of the methods and formulations described herein and/or known in the art.

Where OP-3 comprises part of a tissue or organ preservation solution, any commercially available preservation solution may be used to advantage. For example, useful solutions known in the art include Collins solution, Wisconsin solution, Belzer solution, Eurocollins solution and lactated Ringer's solution. Generally, an organ preservation solution usually possesses one or more of the following properties: (a) an osmotic pressure substantially equal to that of the inside of a mammalian cell,(solutions typically are

hyperosmolar and have K+ and/or Mg++ ions present in an amount sufficient to produce an osmotic pressure slightly higher than the inside of a mammalian cell); (b) the solution typically is capable of maintaining substantially normal ATP levels in the cells; and (c) the solution usually allows optimum maintenance of glucose metabolism in the cells. Organ preservation solutions also may contain anticoagulants, energy sources such as glucose, fructose and other sugars, metabolites, heavy metal chelators, glycerol and other materials of high viscosity to enhance survival at low temperatures, free oxygen radical inhibiting and/or scavenging agents and a pH indicator. A detailed description of preservation solutions and useful components may be found, for example, in US Patent No. 5,002,965.

OP-3 is envisioned to be useful in enhancing viability of any organ or living tissue to be transplanted. The morphogens may be used to particular advantage in lung, heart, liver, kidney or pancreas transplants, as well as in the transplantation and/or grafting of bone marrow, skin, gastrointestinal mucosa, and other living tissues.

As will be appreciated by those skilled in the art, the concentration of the compounds described in a therapeutic composition will vary depending upon a number of factors, including the dosage of the drug to be administered, the chemical characteristics (e.g.,

hydrophobicity) of the compounds employed, and the route of administration. The preferred dosage of drug to be administered also is likely to depend on such variables as the type and extent of tissue loss or defect, the overall health status of the particular patient, the relative biological efficacy of the compound selected, the formulation of the compound, the presence and types of excipients in the formulation, and the route of administration. In general terms, the compounds of this invention may be provided in an aqueous physiological buffer solution containing about 0.001 to 10% w/v compound for parenteral administration. Typical dose ranges are from about 10 ng/kg to about 1 g/kg of body weight per day; a preferred dose range is from about 0.1 μg/kg to 100 mg/kg of body weight. No obvious morphogen-induced pathological lesions are induced when mature morphogen (e.g., OP-1, 20 μg) is administered daily to normal growing rats for 21 consecutive days. Moreover, 10 μg systemic injections of morphogen (e.g., OP-1) injected daily for 10 days into normal newborn mice does not produce any gross abnormalities.

II.B Matrix Preparation

A morphogenically active fragment of OP-3 may be implanted surgically, dispersed in a biocompatible, preferably in vivo biodegradable matrix appropriately modified to provide a structure or scaffold in which the OP-3 may be dispersed and which allows the differentiation and proliferation of migrating

progenitor cells. The matrix also may provide signals capable of directing the tissue specificity of the differentiating cells, as well as providing a morphogenically permissive environment, being essentially free of growth inhibiting signals.

The formulated matrix may be shaped as desired in anticipation of surgery or may be shaped by the physician or technician during surgery. Thus, the material may be used in topical, subcutaneous, intraperitoneal, or intramuscular implants to repair tissue or to induce its growth de novo. The matrix preferably is biodegradable in vivo, being slowly absorbed by the body and replaced by new tissue growth, in the shape or very nearly in the shape of the implant. The matrix also may be particulate in nature.

Details of how to make and how to use the matrices useful in this invention are disclosed below.

II.B(i) Tissue-Derived Matrices

Suitable biocompatible, in vivo biodegradable acellular matrices may be prepared from naturally-occurring tissue. The tissue is treated with suitable agents to substantially extract the cellular, nonstructural components of the tissue. The agents also should be capable of extracting any morphogenesis inhibiting components associated with the tissue. The resulting material is a porous, acellular matrix, substantially depleted in nonstructurally-associated components.

The matrix also may be further treated with agents that modify the matrix, increasing the number of pores and micropits on its surfaces. Those skilled in the art will know how to determine which agents are best suited to the extraction of nonstructural components for different tissues. For example, soft tissues such as liver and lung may be thin-sectioned and exposed to a nonpolar solvent such as, for example, 100% ethanol, to destroy the cellular structure of the tissue and extract nonstructural components. The material then may be dried and pulverized to yield nonadherent porous particles or it may be maintained as a gel-like solution. Structural tissues such as cartilage and dentin where collagen is a primary proteinaceous component may be demineralized and extracted with guanidinium hydrochloride, essentially following the method of Sampath et al. (1983) PNAS 80:6591-6595. For example, pulverized and demineralized dentin is extracted with five volumes of 4M guanidinium-HCl, 50mM Tris-HCl, pH 7.0 for 16 hours at 4°C. The suspension then is filtered. The insoluble material that remains is collected and used to fabricate the matrix. The material is mostly collagenous in matter. It is devoid of morphogenic activity. The matrix particles may further be treated with a collagen fibril-modifying agent that extracts potentially unwanted components from the matrix, and alters the surface structure of the matrix material. Useful agents include acids, organic solvents or heated aqueous media. A detailed description of these matrix treatments are disclosed, for example, in U.S. Patent No. 4,975,526 and PCT publication US90/00912, published September 7, 1990 (WO90/10018).

The currently most preferred agent is a heated aqueous fibril-modifying medium such as water, to increase the matrix particle surface area and porosity. The currently most preferred aqueous medium is an acidic aqueous medium having a pH of less than about 4.5, e.g., within the range of about pH 2 - pH 4 which may help to "swell" the collagen before heating. 0.1% acetic acid, which has a pH of about 3, currently is most preferred. 0.1 M acetic acid also may be used.

Various amounts of delipidated, demineralized guanidine-extracted bone collagen are heated in the aqueous medium (1g matrix/30ml aqueous medium) under constant stirring in a water jacketed glass flask, and maintained at a given temperature for a predetermined period of time. Preferred treatment times are about one hour, although exposure times of between about 0.5 to two hours appear acceptable. The temperature employed is held constant at a temperature within the range of about 37°C to 65°C. The currently preferred heat treatment temperature is within the range of about 45°C to 60°C.

After the heat treatment, the matrix is filtered, washed, lyophilized and used for implant. Where an acidic aqueous medium is used, the matrix also is preferably neutralized prior to washing and lyophilization. A currently preferred neutralization buffer is a 200mM sodium phosphate buffer, pH 7.0. To neutralize the matrix, the matrix preferably first is

allowed to cool following thermal treatment, the acidic aqueous medium (e.g., 0.1% acetic acid) then is removed and replaced with the neutralization buffer and the matrix agitated for about 30 minutes. The neutralization buffer then may be removed and the matrix washed and lyophilized.

Other useful fibril-modifying treatments include acid treatments (e.g., trifluoroacetic acid and hydrogen fluoride) and solvent treatments such as dichloromethane, acetonitrile, isopropanol and chloroform, as well as particular acid/solvent combinations.

After contact with the fibril-modifying agent, the treated matrix may be washed to remove any extracted components, following a form of the procedure set forth below:

1. Suspend matrix preparation in TBS (Tris-buffered saline) 1g/200 ml and stir at 4°C for 2 hrs; or in 6 M urea, 50 mM Tris-HCl, 500 mM NaCl, pH 7.0 (UTBS) or water and stir at room temperature (RT) for 30 minutes (sufficient time to neutralize the pH);

2. Centrifuge and repeat wash step; and

3. Centrifuge; discard supernatant; water wash residue; and then lyophilize.

Alternatively, suitable matrix materials may be obtained commercially. For example, an extracellular matrix extract such as Matrigel™, (Collaborative Research, Inc., Bedford) derived from mouse sarcoma cells, may be used to advantage.

II.B(ii) Synthetic Matrices

In addition to the naturally-derived tissue-specific matrices described above, useful tissue-specific matrices may be formulated synthetically. These porous biocompatible, in vivo biodegradable synthetic matrices are disclosed in PCT publication US91/03603, published December 12, 1991 (WO91/18558). Briefly, the matrix comprises a porous crosslinked structural polymer of biocompatible, biodegradable collagen and appropriate, tissue-specific glycosaminoglycans as tissue-specific cell attachment factors. Collagen derived from a number of sources may be suitable for use in these synthetic matrices, including insoluble collagen, acid-soluble collagen, collagen soluble in neutral or basic aqueous solutions, as well as those collagens which are commercially available.

Glycosaminoglycans (GAGs) or mucopolysaccharides are hexosamine-containing polysaccharides of animal origin that have a tissue specific distribution, and therefore may be used to help determine the tissue

specificity of the morphogen-stimulated differentiating cells. Reaction with the GAGs also provides collagen with another valuable property, i.e., inability to provoke an immune reaction (foreign body reaction) from an animal host.

Chemically, GAGs are made up of residues of hexosamines glycosidically bound and alternating in a more-or-less regular manner with either hexouronic acid or hexose moieties (see, e.g., Dodgson et al. in Carbohydrate Metabolism and its Disorders (Dickens et al., eds.) Vol. 1, Academic Press (1968)). Useful GAGs include hyaluronic acid, heparin, heparin sulfate, chondroitin 6-sulfate, chondroitin 4-sulfate, dermatan sulfate, and keratin sulfate. Other GAGs are suitable for forming the matrix described herein, and those skilled in the art will either know or be able to ascertain other suitable GAGs using no more than routine experimentation. For a more detailed description of mucopolysaccharides, see Aspinall, Polysaccharides, Pergamon Press, Oxford (1970). For example, as disclosed in U.S. Application Serial No. 529,852, chondroitin-6-sulfate can be used where endochondral bone formation is desired. Heparin sulfate, on the other hand, may be used to formulate synthetic matrices for use in lung tissue repair.

Collagen can be reacted with a GAG in aqueous acidic solutions, preferably in diluted acetic acid solutions. By adding the GAG dropwise into the aqueous

collagen dispersion, coprecipitates of tangled collagen fibrils coated with GAG results. This tangled mass of fibers then can be homogenized to form a homogeneous dispersion of fine fibers and then filtered and dried.

Insolubility of the collagen-GAG products can be raised to the desired degree by covalently cross-linking these materials, which also serves to raise the resistance to resorption of these materials. In general, any covalent cross-linking method suitable for cross-linking collagen also is suitable for cross-linking these composite materials, although crosslinking by a dehydrothermal process is preferred.

When dry, the crosslinked particles are essentially spherical, with diameters of about 500 μm. Scanning electron miscroscopy shows pores of about 20 μm on the surface and 40 μm on the interior. The interior is made up of both fibrous and sheet-like structures, providing surfaces for cell attachment. The voids interconnect, providing access to the cells throughout the interior of the particle. The material appears to be roughly 99.5% void volume, making the material very efficient in terms of the potential cell mass that can be grown per gram of microcarrier.

Another useful synthetic matrix is one formulated from biocompatible, in vivo biodegradable synthetic polymers, such as those composed of glycolic acid, lactic acid and/or butyric acid, including copolymers

and derivatives thereof. These polymers are well described in the art and are available commercially. For example, polymers composed of polyactic acid (e.g., MW 100 kDa), 80% polylactide/20% glycoside or poly 3-hydroxybutyric acid (e.g., MW 30 kDa) all may be purchased from PolySciences, Inc. The polymer compositions generally are obtained in particulate form. In addition, one can alter the morphology of the polymer compositions, for example to increase porosity, using any of a number of particular solvent treatments known in the art. Where the morphogen is adsorbed to the matrix surface, the steps preferably are performed under conditions which avoid hydrolysis of the polymers (e.g., non-aqueous conditions such as in an ethanol-trifluoro-acetic acid solution).

The OP-3 proteins described herein can be combined and dispersed in a suitable matrix using any of the methods described below:

1. Ethanol Precipitation

Matrix is added to the morphogen dissolved in guanidine-HCl. Samples are vortexed and incubated at a low temperature. Samples are then further vortexed. Cold absolute ethanol is added to the mixture which is then stirred and incubated. After centrifugation (microfuge, high speed) the supernatant is discarded. The matrix is washed with cold concentrated ethanol in water and then lyophilized.

2. Acetonitrile Trifluoroacetic Acid Lyophilization

In this procedure, a morphogenically active fragment of OP-3 in an acetonitrile trifluroacetic acid (ACN/TFA) solution is added to the carrier material. Samples are vigorously vortexed many times and then lyophilized.

3. Buffered Saline Lyophilization

A preparation of a morphogenically active fragment of OP-3 in physiological saline also may be vortexed with the matrix and lyophilized to produce morphogenically active material.

Tissue morphogenesis requires a morphogenically permissive environment. Clearly, in fully-functioning healthy tissue that is not composed of a permanently renewing cell population, there must exist signals to prevent continued tissue growth. Thus, it is postulated that there exists a control mechanism, such as a feedback control mechanism, which regulates the control of cell growth and differentiation. In fact, it is known that both TGF-β, and MIS are capable of inhibiting cell growth when present at appropriate concentrations. In addition, using the bone model system it can be shown that osteogenic devices comprising a bone-derived carrier (matrix) that has been demineralized and guanidine-extracted to substantially remove the noncollagenous proteins does

allow endochondral bone formation when implanted in association with an osteoinductive morphogen. If, however, the bone-derived carrier is not demineralized but rather is washed only in low salt, for example, induction of endochondral bone formation is inhibited, suggesting the presence of one or more inhibiting factors within the carrier.

III. Examples

Example 1. Recombinant Production of OP-3

OP-3 proteins useful in the methods and compositions of this invention may be purified from natural sources or produced using standard recombinant methodology. General considerations for the recombinant production of OP3 morphogens are described below.

A. Identification of Novel mOP-3 Sequences

A genetic sequence encoding the morphogenic OP-3 protein was identified using a 0.3 kb EcoRI-BamH1 OP-2 fragment from a mouse OP-2 cDNA as a hybridization probe, specific to the mid-pro region of OP-2 (corresponding to amino acid residues 125 to 225 of the pre-pro protein) essentially as described in USSN 667,274. The ^{32}P-labeled probe was prepared using the random hexanucleotide priming method, and the

hybridizations were performed using the following conditions: 40% formamide, 5 X SSPE, 5 X Denhardt's Solution, 0.1% SDS, at 37°C overnight, and washing in 0.1 X SSPE, 0.1% SDS at 50°C. Approximately 1 X 10^6 phages from a mouse cDNA (carried in lambda zapII) library made from the teratocarcinoma cell line PCC4 (Stratagene, Inc., La Jolla, CA, cat # 936301) were screened. This screening yielded four individual clones which were purified over three rounds of screening. The plasmid DNA containing the cDNAs was obtained using the lambda zapII excision process following manufacturer's directions. Three of the four clones were shown by DNA sequencing to encode OP-3. The DNA sequence, referred to herein as mOP-3 and described in Seq. ID No. 1, was identified by this procedure.

The isolated mOP-3 DNA sequence, in accordance with other known morphogens, encodes a protein comprising a "pro" region (defined essentially by residues 20-260 or 20-263 of Seq. ID No. 1) and a mature region (defined essentially by residues 261-399 or 264-399 of Seq. ID No. 1), including a functional domain comprising the conserved cysteine skeleton.

Like OP-2, OP-3 is marked by an eighth cysteine within the seven cysteine domain (e.g., at position 338 of Seq. ID No.1). The extra cysteine likely helps stabilize the folded structure, possibly by providing inter-molecular disulfide bonding. The extra cysteine

also allows for heterodimer formation between OP-3 and another morphogen comprising the "eighth" cysteine, like OP-2 for example, or a modified OP-1, wherein an extra cysteine has been inserted at the appropriate location. The extra cysteine also may allow tetramer formation. The extra cysteine does not inhibit synthesis or reduce the stability of the translated sequence significantly as expressed proteins comprising the extra cysteine are readily detected by SDS gel electrophoresis. A primary glycosylation site occurs just C terminal to the extra cysteine in both OP-2 and OP-3, which may provide a protective effect.

The cDNA sequences for both human and mouse OP-2 are provided in Seq. ID Nos. 7 and 9, and the genomic sequence for human OP-2 is provided in Seq. ID No. 11, wherein the exons defining the coding region of these proteins are indicated. The exon boundaries also are indicated in Fig. 1, described below. The human OP-2 locus was isolated from a genomic library (Clontech, EMBL-3 #HL1067J) on three overlapping phage clones, using standard cloning procedures. The OP-2 coding information was spread over 27 kb and, like OP-1, contains 7 exons. A comparison of exon-intron boundaries in the 7 cysteine domain showed matching locations with those of OP-1. The first OP-2 exon contains 334 bp of coding sequence (111 amino acids), including the signal peptide, and is followed by the largest intron (14.6 kb). The second exon (190 bp, 64 amino acids) is separated by a short intron (0.4 kb)

from exon 3 (149 bp, 49 amino acids). It follows a large third intron of 9.5 kb. The fourth exon (195 bp, 65 amino acids) encodes the maturation site ("OP-2-Ala") and is followed by a 0.8 kb intron. The 7 cysteine domains resides on exons 5 to 7: exon 5 (80 bp, 27 amino acids) encodes the first cysteine of mature OP-2 and is followed by intron 5 (0.5 kb in length), exon 6 (111 bp, 37 amino acids) is separated by a 2.5 kb intron from the seventh, last exon with 147 bp (49 amino acids) of coding sequence. As stated above, the exon-intron boundaries are conserved between human OP-1 and OP-2, two different members of the morphogen family of proteins. By analogy, the exon-intron boundaries between human and mouse OP-2, two species variants of a morphogen, are anticipated to be conserved as well.

Figure 1 shows the alignment of the murine OP-2 and murine OP-3 coding regions of the cDNA. The exon boundaries are indicated by bars beneath the sequence. Both sequences have the same number of nucleotides. The nucleotide sequence is about 80% conserved in the N-terminal and C-terminal regions. In the figure, nucleotide identity between the sequences is indicated by stippling. In addition, the central region of the sequence is highly conserved and this conserved region falls into the boundaries of exon 2 and 3. There are only three nucleotide differences in this region, indicated in the figure by diamonds.

The high degree of conservation in the nucleotide sequences indicates that OP-2 and OP-3 likely share the nucleotide sequence of exon 2 and 3. The different proteins may result from alternatively spliced transcripts, or they may arise from independent genes which share part of their coding sequence. Intron 1, which lies upstream of exon 2 in OP-2 (see Seq. ID No.11) is large (14.6kb) and could include the start of the OP-3 gene and/or its first exon sequence. Certainly, as has been found for other mammalian genes, one or more of the introns of these morphogens may include sequences having a transcription regulatory function.

Using the screening procedure described herein and in USSN 752,764, and the labelled OP-2 fragment, or preferably a labelled OP-3 fragment, OP-3 genetic sequences from other species and other libraries may be isolated. Alternatively, or in addition, a probe to the N-terminal region of the mature protein, or the 3' noncoding region flanking and immediately following the stop codon, also may be used to screen for other OP-3 species variants. These sequences vary substantially among the morphogens and represent morphogen-specific sequences. Mammalian cell expression of OP-3 readily can be achieved using COS (simian kidney ATCC, CRL-1650) or CHO (Chinese hamster ovary) cells (e.g., CHO-DXBII, from Lawrence Chasin, Columbia University, NY). An exemplary protocol for mammalian cell expression is provided below. Other useful eukaryotic cell systems include the insect/baculovirus system or the mammalian complement system.

B. Expression of Novel OP-3 Sequences

To express the OP-3 protein, the OP-3 DNA is subcloned into an insertion site of a suitable, commercially available pUC-type vector (e.g., pUC-19, ATCC #37254, Rockville, MD), along with a suitable promoter/enhancer sequences and 3' termination sequences. Currently preferred promoter/enhancer sequences are the CMV-MIE promoter (human cytomegalovirus major intermediate-early promoter, preferably the intron-free or "short" form of the promoter) and the mouse mammary tumor virus promoter (mMTV) boosted by the rous sarcoma virus LTR enhancer sequence (e.g., from Clontech, Inc., Palo Alto). Expression also may be further enhanced using transactivating enhancer sequences. The plasmid also preferably contains a selectable marker, most preferably an amplifiable marker such as DHFR, e.g., under SV40 early promoter control (ATCC #37148). Transfection, cell culturing, gene amplification and protein expression conditions are standard conditions, well known in the art, such as are described, for example in Ausubel et al., ed., Current Protocols in Molecular Biology, John Wiley & Sons, NY (1989). Briefly, transfected cells are cultured in medium containing 0.1-0.5% dialyzed fetal calf serum (FCS), stably transfected high expression cell lines obtained by subcloning and evaluated by standard Northern blot. Southern blots also are used to assess the state of integrated OP-3 sequences and the extent of their copy number amplification.

Chimeric OP-3 morphogens, e.g., comprising an OP-3 active domain and, for example, part or all of a pro domain from another, different morphogen may be constructed using standard recombinant DNA technology and/or an automated DNA synthesizer to construct the desired sequence. Useful chimeras include those wherein the non-OP-3 sequence is joined to the OP-3 sequence encoding the mature OP-3 protein, and the non-OP-3 sequence encodes part or all of the sequence between the signal peptide processing site and the "Arg-Xaa-Xaa-Arg" processing sequence from at least one morphogen. Alternatively, the non-OP-3 sequence may be joined to an OP-3 sequence encoding, for example, the 6 or 7 cysteine skeletons, wherein the non-OP-3 sequence includes the sequence encoding the N-terminus of the mature protein. As will be appreciated by persons skilled in the art, the non-OP-3 sequences may be composed of sequences from one or morphogens and/or may comprise novel biosynthetic sequences.

Mammalian expression of a biosynthetic gene construct encoding a chimeric OP1-OP3 polypeptide chain is demonstrated in the immunoblot presented in Fig. 2. A vector carrying the construct under CMV promoter control was transfected into CHO cells (CHO-DXB11) using standard procedures and as described herein.

A chimeric gene was constructed by replacing the conserved seven cysteine domain of OP-1 with that of OP-3. The resulting chimeric gene contains the entire pre-pro-domain of human OP-1 and the region of mature

OP-1 between the maturation site and the first cysteine of the conserved C-terminal seven cysteine domain, fused to the conserved seven cysteine domain of mouse OP-3, but with two arginine residues in place of the native lysine residues found in OP-3 at the start of the seven cysteine domain.

The gene fusion was accomplished by splicing the SacI site of OP-3 (near the first cysteine of the seven cysteine domain) with a newly created SacI site in OP-1, created at the matching residues by silent mutagenesis. The SacI site encodes the Glu-Leu dipeptide in the sequence Cys-Arg-Arg-His-Glu-Leu of OP-1 and Cys-Lys-Lys-His-Glu-Leu of OP-3, respectively.

The chimeric gene was placed downstream of the CMV (Cytomegalovirus) MIE "short" (intron-free) promoter and upstream of the SV40 transcriptional terminator in a pUC vector. This plasmid was cotransfected with DNA encoding the DHFR marker and viral trans-activating elements (e.g., VA1, E1A) into a CHO dhfr(-) host and subjected to Methotrexate selection and one round of amplification at 1 mM Methotrexate including subcloning. 10 μl of "spent" culture supernatant (3 days old) was analysed by "Western blot" (immunoblot), as follows.

The 10 μl harvested medium was briefly heated with concentrated SDS sample buffer, containing β-mercapto ethanol (5%) and directly analysed by electrophoresis on a 15% SDS- polyacrylamide gel (in the buffer system

of Laemmli) along with a set of prestained molecular weight standards (Bio-rad, Richmond, CA). Proteins were transferred from the gel to Immobilon membrane by the "Western blot" procedure. The chimeric OP-1/OP-3 protein was detected by reaction with rabbit serum raised against a synthetic peptide representing the first 17 amino acids of mature OP-1, starting with serine-threonine-glycine-serine-. Authentic recombinant OP-1, expressed in CHO cells was included for comparison. In the figure sample lanes were as follows: lane 1: OP-1; lanes 4, 5, 6, 7, and 8: chimeric OP-1/OP-3; lanes 9 and 10: prestained molecular weight standards. The apparent mobility of the recombinant proteins, at approximately 20 kDa on this gel, is due to glycosylation of the OP-1 and OP-3 proteins which may also be the cause of the multiple species observed.

The expressed protein then can be purified as follows. For a typical 2L preparation of transfected mammalian cells conditioned in 0.5% FCS, for example, the total protein is typically about 700 mg. The amount of OP-3 in the media, estimated by Western blot, is between about 0.1-5.0 mg. OP-3 media then is diluted in a low salt, physiologically buffered 6M urea solution, and loaded onto an S-Sepharose column, which acts as a strong cation exchanger. OP-3 binds to the column in low salt, and serum proteins are removed. The column subsequently is developed with an NaCl gradient, e.g., 0.1M NaCl-1.0M NaCl, in 6M urea, 20mM HEPES, pH 7.0. Most contaminants are removed at the start of the gradient, and OP-3 is eluted primarily at a higher salt concentration.

The sample then is loaded onto a phenyl-Sepharose column (hydrophobic interaction chromatography). OP-3 binds phenyl-Sepharose in the presence of high concentrations of a weak chaotropic salt (e.g., 1M $(NH_4)_2SO_4$ in a physiologically buffered 6M urea solution). Once OP-3 is bound, the column is developed with a decreasing ammonium sulfate gradient, e.g., 0.6M-0.0M $(NH_4)_2SO_4$ gradient in a physiologically buffered, 6M urea solution. Again, most contaminants are removed at the start of the gradient, and OP-3 elutes primarily at low or no ammonium sulfate concentrations.

The OP-3 eluted from the phenyl-Sepharose column then is dialyzed against water, and prepared for loading onto a reverse phase chromatography column (e.g., C-18 HPLC), for example, by dialyzing against 30% acetonitrile, 0.1% TFA.

An alternative chromatography protocol is to perform the S-Sepharose chromatography in the absence of 6 M urea. The bound proteins then are eluted with salt step elutions (e.g., 0.1-0.6M NaCl). Remaining OP-3 then can be eluted in the presence of 6M urea. The 6M urea elution also may be used in place of the non-urea elution to achieve maximum recovery in one step. In addition, OP-3 may be eluted from the phenyl-Sepharose column in 38% ethanol-0.01% TFA, thereby eliminating the need to dialyze the eluent before applying it to the C-18 column. Finally, multiple C-18 columns may be used (e.g., three), to further enhance purification and concentration of the protein.

OP-3 also will bind hydroxyapatite efficiently, typically in the absence of 6 M urea and at low phosphate concentrations (less than 5 mM phosphate). Bound OP-3 can be removed from the column with an elution gradient of about .001-0.5M step elution of phosphate in a physiologically buffered solution. Additionally, urea (6M) may be added during the elution step.

Other related chromatography methods also may be useful in purifying OP-3 from eucaryotic cell culture systems. For example, heparin-Sepharose may be used in combination with the S-Sepharose column. Alternatively, immobilized metal-ion affinity chromatography (IMAC) (e.g., Cu^{2+} or Zn^{+}) and a physiologically buffered phosphate solution may be used to advantage.

C. Soluble OP3 Complexes

A currently preferred form of the OP-3 morphogen useful in therapeutic formulations, having improved solubility in aqueous solutions and consisting essentially of amino acids, is a dimeric morphogenic protein comprising at least the 100 amino acid peptide sequence having the pattern of seven or more cysteine residues characteristic of the morphogen family complexed with a peptide comprising part or all of a pro region of a member of the morphogen family, or an

allelic, species or other sequence variant thereof. Preferably, the dimeric morphogenic protein is complexed with two peptides. Also, the dimeric morphogenic protein preferably is noncovalently complexed with the pro region peptide or peptides. The pro region peptides also preferably comprise at least the N-terminal eighteen amino acids that define the OP-3 morphogen pro region (e.g., residues 18-35 of Seq. ID No. 1). In a most preferred embodiment, peptides defining substantially the full length pro region are used.

Other soluble forms of morphogens include dimers of the uncleaved pro forms of these proteins, as well as "hemi-dimers" wherein one subunit of the dimer is an uncleaved pro form of the protein, and the other subunit comprises the mature form of the protein, including truncated forms thereof, preferably noncovalently associated with a cleaved pro domain peptide.

As described above, useful pro domains include the full length pro regions, as well as various truncated forms hereof, particularly truncated forms cleaved at proteolytic Arg-Xaa-Xaa-Arg cleavage sites. In OP-3, possible pro sequences cleaved at Arg-Xaa-Xaa-Arg sites include sequences defined by residues 18-260 of Seq. ID No. 1 (anticipated full length form); or by residues 18-263. Accordingly, currently preferred pro sequences are those encoding the full length form of the pro region for OP-3 or another, known morphogen. Other pro

sequences contemplated to have utility include biosynthetic pro sequences, particularly those that incorporate a sequence derived from the N-terminal portion of one or more morphogen pro sequences.

As will be appreciated by those having ordinary skill in the art, useful sequences encoding the pro region may be obtained from genetic sequences encoding known morphogens. Alternatively, chimeric pro regions can be constructed from the sequences of one or more known morphogens. Still another option is to create a synthetic sequence variant of one or more known pro region sequences.

In another preferred aspect, useful pro region peptides include polypeptide chains comprising an amino acid sequence encoded by a nucleic acid that hybridizes under stringent conditions with a DNA or RNA sequence encoding at least the N-terminal eighteen amino acids of the pro region sequence for OP-3 e.g., nucleotides 120-173 of Seq. ID No. 1.

In yet another preferred aspect, useful pro region peptides include polypeptide chains comprising an amino acid sequence encoded by a nucleic acid that hybridizes under stringent conditions with a DNA or RNA sequence encoding at least the N-terminal eighteen amino acids of the pro region sequence for OP1 or OP2, e.g., nucleotides 136-192 and 152-211 of Seq. ID No. 3 and 7, respectively.

C.1. Isolation of Soluble morphogen complex from conditioned media or body fluid

Morphogens are expressed from mammalian cells as soluble complexes. Typically, however the complex is disassociated during purification, generally by exposure to denaturants often added to the purification solutions, such as detergents, alcohols, organic solvents, chaotropic agents and compounds added to reduce the pH of the solution. Provided below is a currently preferred protocol for purifying the soluble proteins from conditioned media (or, optionally, a body fluid such as serum, cerebro-spinal or peritoneal fluid), under non-denaturing conditions. The method is rapid, reproducible and yields isolated soluble morphogen complexes in substantially pure form.

Soluble OP-3 morphogen complexes can be isolated from conditioned media using a simple, three step chromatographic protocol performed in the absence of denaturants. The protocol involves running the media (or body fluid) over an affinity column, followed by ion exchange and gel filtration chromatographies. The affinity column described below is a Zn-IMAC column. The present protocol has general applicability to the purification of a variety of morphogens, all of which are anticipated to be isolatable using only minor modifications of the protocol described below. An alternative protocol also envisioned to have utility an immunoaffinity column, created using standard procedures and, for example, using antibody specific

for a the OP-3 pro domain (complexed, for example, to a protein A-conjugated Sepharose column.) Protocols for developing immunoaffinity columns are well described in the art, (see, for example, Guide to Protein Purification, M. Deutscher, ed., Academic Press, San Diego, 1990, particularly sections VII and XI.)

In this experiment OP-1 was expressed in mammalian CHO (chinese hamster ovary) cells as described in the art (see, for example, international application US90/05903 (WO91/05802).) The CHO cell conditioned media containing 0.5% FBS was initially purified using Immobilized Metal-Ion Affinity Chromatography (IMAC). The soluble OP-1 complex from conditioned media binds very selectively to the Zn-IMAC resin and a high concentration of imidazole (50 mM imidazole, pH 8.0) is required for the effective elution of the bound complex. The Zn-IMAC step separates the soluble OP-1 from the bulk of the contaminating serum proteins that elute in the flow through and 35 mM imidazole wash fractions. The Zn-IMAC purified soluble OP-1 is next applied to an S-Sepharose cation-exchange column equilibrated in 20 mM $NaPO_4$ (pH 7.0) with 50 mM NaCl. This S-Sepharose step serves to further purify and concentrate the soluble OP-1 complex in preparation for the following gel filtration step. The protein was applied to a Sephacryl S-200HR column equilibrated in TBS. Using substantially the same protocol, soluble morphogens also may be isolated from one or more body fluids, including serum, cerebro-spinal fluid or peritoneal fluid.

IMAC was performed using Chelating-Sepharose (Pharmacia) that had been charged with three column volumes of 0.2 M $ZnSO_4$. The conditioned media was titrated to pH 7.0 and applied directly to the ZN-IMAC resin equilibrated in 20 mM HEPES (pH 7.0) with 500 mM NaCl. The Zn-IMAC resin was loaded with 80 mL of starting conditioned media per mL of resin. After loading, the column was washed with equilibration buffer and most of the contaminating proteins were eluted with 35 mM imidazole (pH 7.0) in equilibration buffer. The soluble OP-1 complex then is eluted with 50 mM imidazole (pH 8.0) in 20 mM HEPES and 500 mM NaCl.

The 50 mM imidazole eluate containing the soluble OP-1 complex was diluted with nine volumes of 20 mM $NaPO_4$ (pH 7.0) and applied to an S-Sepharose (Pharmacia) column equilibrated in 20 mM $NaPO_4$ (pH 7.0) with 50 mM NaCl. The S-Sepharose resin was loaded with an equivalent of 800 mL of starting conditioned media per mL of resin. After loading the S-Sepharose column was washed with equilibration buffer and eluted with 100 mM NaCl followed by 300 mM and 500 mM NaCl in 20 mM $NaPO_4$ (pH 7.0). The 300 mM NaCl pool was further purified using gel filtration chromatography. Fifty mls of the 300 mm NaCl eluate was applied to a 5.0 X 90 cm Sephacryl S-200HR (Pharmacia) equilibrated in Tris buffered saline (TBS), 50 mM Tris, 150 mM NaCl (pH 7.4). The column was eluted at a flow rate of 5 mL/minute collecting 10 mL fractions. The apparent molecular of the soluble OP-1 was determined by comparison to protein molecular weight standards

(alcohol dehydrogenase (ADH, 150 kDa), bovine serum albumin (BSA, 68 kDa), carbonic anhydrase (CA, 30 kDa) and cytochrome C (cyt C, 12.5 kDa). The purity of the S-200 column fractions was determined by separation on standard 15% polyacrylamide SDS gels stained with coomassie blue. The identity of the mature OP-1 and the pro-domain was determined by N-terminal sequence analysis after separation of the mature OP-1 from the pro-domain using standard reverse phase C18 HPLC.

The soluble OP-1 complex elutes with an apparent molecular weight of 110 kDa. This agrees well with the predicted composition of the soluble OP-1 complex with one mature OP-1 dimer (35-36 kDa) associated with two pro-domains (39 kDa each). Purity of the final complex can be verified by running the appropriate fraction in a reduced 15% polyacrylamide gel.

The complex components can be verified by running the complex-containing fraction from the S-200 or S-200HR columns over a reverse phase C18 HPLC column and eluting in an acetonitrile gradient (in 0.1% TFA), using standard procedures. The complex is dissociated by this step, and the pro domain and mature species elute as separate species. These separate species then can be subjected to N-terminal sequencing using standard procedures (see, for example, Guide to Protein Purification, M. Deutscher, ed., Academic Press, San Diego, 1990, particularly pp. 602-613), and the identity of the isolated 36kD, 39kDa proteins confirmed as mature morphogen and isolated, cleaved pro domain, respectively. N-terminal sequencing of the

isolated pro domain from mammalian cell produced OP-1 revealed 2 forms of the pro region, the intact form (beginning at residue 30 of Seq. ID No. 16) and a truncated form, (beginning at residue 48 of Seq. ID No. 16.) N-terminal sequencing of the polypeptide subunit of the isolated mature species reveals a range of N-termini for the mature sequence, beginning at residues 293, 300, 313, 315, 316, and 318, of Seq. ID No. 16, all of which are active as demonstrated by the standard bone induction assay.

C.2. In Vitro Soluble Morphogen Complex Formation

As an alternative to purifying soluble complexes from culture media or a body fluid, soluble complexes may be formulated from purified pro domains and mature dimeric species. Successful complex formation apparently requires association of the components under denaturing conditions sufficient to relax the folded structure of these molecules, without affecting disulfide bonds. Preferably, the denaturing conditions mimic the environment of an intracellular vesicle sufficiently such that the cleaved pro domain has an opportunity to associate with the mature dimeric species under relaxed folding conditions. The concentration of denaturant in the solution then is decreased in a controlled, preferably step-wise manner, so as to allow proper refolding of the dimer and pro regions while maintaining the association of the pro domain with the dimer. Useful denaturants include 4-6M urea or guanidine hydrochloride (GuHCl), in buffered solutions of pH 4-10, preferably pH 6-8. The soluble complex then is formed by controlled dialysis or dilution into a solution having a final denaturant

concentration of less than 0.1-2M urea or GuHCl, preferably 1-2 M urea of GuHCl, which then preferably can be diluted into a physiological buffer. Protein purification/renaturing procedures and considerations are well described in the art, and details for developing a suitable renaturing protocol readily can be determined by one having ordinary skill in the art. One useful text one the subject is Guide to Protein Purification, M. Deutscher, ed., Academic Press, San Diego, 1990, particularly section V. Complex formation also may be aided by addition of one or more chaperone proteins.

C3. Stability of Soluble Morphogen Complexes

The stability of the highly purified soluble morphogen complex in a physiological buffer, e.g., tris-buffered saline (TBS) and phosphate-buffered saline (PBS), can be enhanced by any of a number of means. Currently preferred is by means of a pro region that comprises at least the first 18 amino acids of the pro sequence (e.g., residues 18-35 of Seq. ID NO. 1 for OP-3), and preferably is the full length pro region. Residues 18-35 show sequence homology to the N-terminal portion of other morphogens and are believed to have particular utility in enhancing complex stability for all morphogens. Other useful means for enhancing the stability of soluble morphogen complexes include three classes of additives. These additives include basic amino acids (e.g., L-arginine, lysine and betaine); nonionic detergents (e.g., Tween 80 or NonIdet P-120); and carrier proteins (e.g., serum albumin and casein).

Useful concentrations of these additives include 1-100 mM, preferably 10-70 mM, including 50 mM, basic amino acid;, 0.01-1.0%, preferably 0.05-0.2%, including 0.1% (v/v) nonionic detergent;, and 0.01-1.0%, preferably 0.05-0.2%, including 0.1% (w/v) carrier protein.

Example 2. Mitogenic Effect of OP-3

2.1 Mitogenic Effect of Morphogen on Rat and Human Osteoblasts

The following example can be used to demonstrate the ability of OP-3 to induce proliferation of osteoblasts in vitro using the following assay. In this and all examples involving osteoblast cultures, rat osteoblast-enriched primary cultures preferably are used. Although these cultures are heterogeneous in that the individual cells are at different stages of differentiation, the culture is believed to more accurately reflect the metabolism and function of osteoblasts in vivo than osteoblast cultures obtained from established cell lines. Unless otherwise indicated, all chemicals referenced are standard, commercially available reagents, readily available from a number of sources, including Sigma Chemical, Co., St. Louis; Calbiochem, Corp., San Diego and Aldrich Chemical Co., Milwaukee.

Rat osteoblast-enriched primary cultures are prepared by sequential collagenase digestion of newborn suture-free rat calvaria (e.g., from 1-2 day-old animals, Long-Evans strain, Charles River Laboratories, Wilmington, MA), following standard procedures, such as are described, for example, in Wong et al., (1975) PNAS

72:3167-3171. Rat osteoblast single cell suspensions then are plated onto a multi-well plate (e.g., a 24 well plate) at a concentration of 50,000 osteoblasts per well in alpha MEM (modified Eagle's medium, Gibco, Inc., Long Island) containing 10% FBS (fetal bovine serum), L-glutamine and penicillin/streptomycin. The cells are incubated for 24 hours at 37°C, at which time the growth medium is replaced with alpha MEM containing 1% FBS and the cells incubated for an additional 24 hours so that cells are in serum-deprived growth medium at the time of the experiment.

The cultured cells are divided into three groups: (1) wells which receive, for example, 0.1, 1.0, 10.0, 40 and 80.0 ng of OP-3; (2) wells which receive 0.1, 1.0, 10.0 and 40 ng of a local-acting growth factor (e.g., TGF-β); and (3) the control group, which receive no growth factors. The cells then are incubated for an additional 18 hours after which the wells are pulsed with 2μCi/well of ^{3}H-thymidine and incubated for six more hours. The excess label then is washed off with a cold solution of 0.15 M NaCl and then 250 μl of 10% tricholoracetic acid is added to each well and the wells incubated at room temperature for 30 minutes. The cells then are washed three times with cold distilled water, and lysed by the addition of 250 μl of 1% sodium dodecyl sulfate (SDS) for a period of 30 minutes at 37°C. The resulting cell lysates are harvested using standard means well known in the art, and the incorporation of ^{3}H-thymidine into cellular DNA determined by liquid scintillation as an indication of mitogenic activity of the cells. In the experiment, OP-3 is anticipated to stimulate ^{3}H-thymidine incorporation into DNA, and thus promote osteoblast

cell proliferation. By contrast, the effect of TGF-β is transient and biphasic. At high concentrations, TGF-β has no significant effect on osteoblast cell proliferation.

The in vitro effect of OP-3 on osteoblast proliferation also may be evaluated using human primary osteoblasts (obtained from bone tissue of a normal adult patient and prepared as described above) and on human osteosarcoma-derived cell lines. In all cases OP-3 is anticipated to induce cell proliferation in accordance with the morphogen's ability to induce endochondral bone formation (see Example 7, below).

2.2 Progenitor Cell Stimulation

The following example demonstrates the ability of OP-3 to stimulate the proliferation of mesenchymal progenitor cells. Useful naive stem cells include pluripotential stem cells, which may be isolated from bone marrow or umbilical cord blood using conventional methodologies, (see, for example, Faradji et al., (1988) Vox Sang., 55 (3):133-138 or Broxmeyer et al., (1989) PNAS 86:3828-3832), as well as naive stem cells obtained from blood. Alternatively, embryonic cells (e.g., from a cultured mesodermal cell line) may be useful.

Another method for obtaining progenitor cells and for determining the ability of OP-3 fragments to stimulate cell proliferation is to capture progenitor cells from an in vivo source. For example, a biocompatible matrix material able to allow the influx of migratory progenitor cells may be implanted at an in

vivo site long enough to allow the influx of migratory progenitor cells. For example, a bone-derived, guanidine-extracted matrix, formulated as disclosed for example in Sampath et al. ((1983) PNAS 80:6591-6595), or U.S. Patent No. 4,975,526, may be implanted into a rat at a subcutaneous site, essentially following the method of Sampath et al. After three days the implant is removed, and the progenitor cells associated with the matrix dispersed and cultured.

Progenitor cells, however obtained, then are incubated in vitro with OP-3 under standard cell culture conditions well described in the art and described hereinabove. In the absence of external stimuli, the progenitor cells do not, or only minimally, proliferate on their own in culture. However, progenitor cells cultured in the presence of a morphogenically active fragment of OP-3 are anticipated to proliferate. Cell growth can be determined visually or spectrophotometrically using standard methods well known in the art.

Example 3. Morphogen-Induced Cell Differentiation

3.1 Embryonic Mesenchyme Differentiation

Morphogenically active fragments of OP-3 can be utilized to induce cell differentiation. The ability of OP-3 to induce cell differentiation can be demonstrated by culturing early mesenchymal cells in the presence of OP-3 and then studying the histology of

the cultured cells by staining with toluidine blue using standard cell culturing and cell staining methodologies well described in the art. For example, it is known that rat mesenchymal cells destined to become mandibular bone, when separated from the overlying epithelial cells at stage 11 and cultured in vitro under standard tissue culture conditions, e.g., in a chemically defined, serum-free medium, containing for example, 67% DMEM (Dulbecco's modified Eagle's medium), 22% F-12 medium, 10mM Hepes pH 7, 2mM glutamine, 50 μg/ml transferrin, 25 μg/ml insulin, trace elements, 2mg/ml bovine serum albumin coupled to oleic acid, with HAT (0.1 mM hypoxanthine, 10μM aminopterin, 12 μM thymidine, will not continue to differentiate. However, if these same cells are left in contact with the overlying endoderm for an additional day, at which time they become stage 12 cells, they will continue to differentiate on their own in vitro to form chondrocytes. Further differentiation into osteoblasts and, ultimately, mandibular bone, requires an appropriate local environment, e.g., a vascularized environment.

Stage 11 mesenchymal cells, cultured in vitro in the presence of OP-3, e.g., 10-100 ng/ml, are anticipated to continue to differentiate in vitro to form chondrocytes just as they continue to differentiate in vitro if they are cultured with the cell products harvested from the overlying endodermal cells. This experiment may be performed with different mesenchymal cells to demonstrate the cell differentiation capability of OP-3 in different tissues.

As another example of morphogen-induced cell differentiation, the ability of OP-3 to induce osteoblast differentiation may be demonstrated in vitro using primary osteoblast cultures, or osteoblast-like cells lines, and assaying for a variety of bone cell markers that are specific markers for the differentiated osteoblast phenotype, e.g., alkaline phosphatase activity, parathyroid hormone-mediated cyclic AMP (cAMP) production, osteocalcin synthesis, and enhanced mineralization rates.

3.2 Alkaline Phosphatase Induction of Osteoblasts by OP-3

The cultured cells in serum-free medium are incubated with, a range of OP-3 concentrations, for example, 0.1, 1.0, 10.0, 40.0 or 80.0 ng OP-3/ml medium; or with a similar range of TGF-β concentrations. 72 hours after the incubation period the cell layer is extracted with 0.5 ml of 1% Triton X-100. The resultant cell extract then, is centrifuged, and 100 µl of the extract is added to 90 µl of paranitrosophenylphospate (PNPP)/glycerine mixture and incubated for 30 minutes in a 37°C water bath and the reaction stopped with 100 µl NaOH. The samples then are run through a plate reader (e.g., Dynatech MR700 plate reader, and absorbance measured at 400 nm, using p-nitrophenol as a standard) to determine the presence and amount of alkaline phosphate activity. Protein concentrations are determined by the Biorad method. Alkaline phosphatase activity is calculated in units/µg protein, where 1 unit=1 nmol p-nitrophenol liberated/30 minutes at 37°C.

OP-3 alone stimulates the production of alkaline phosphatase in osteoblasts, and thus promotes the growth and expression of the osteoblast differentiated phenotype.

The long term effect of OP-3 morphogen on the production of alkaline phosphatase by rat osteoblasts also may be demonstrated as follows.

Rat osteoblasts are prepared and cultured in multi-well plates as described above. In this example six sets of 24 well plates are plated with 50,000 rat osteoblasts per well. The wells in each plate, prepared as described above, then are divided into three groups: (1) those which receive, for example, 1 ng of OP-3 per ml of medium; (2) those which receive 40 ng of OP-3 per ml of medium; and (3) those which received 80 ng of OP-3 per ml of medium. Each plate then is incubated for different lengths of time: 0 hours (control time), 24 hours, 48 hours, 96 hours, 120 hours and 144 hours. After each incubation period, the cell layer is extracted with 0.5 ml of 1% Triton X-100. The resultant cell extract is centrifuged, and alkaline phosphatase activity determined as for Example 3.1, using paranitroso-phenylphosphate (PNPP). OP-3 stimulates the production of alkaline phosphatase in osteoblasts in dose-dependent manner so that increasing doses of OP-3 further increase the level of alkaline phosphatase production, and moreover, the OP-3-stimulated elevated levels of alkaline phosphatase in the treated osteoblasts is anticipated to last for an extended period of time.

3.3 OP-3 Protein Induction of PTH-Mediated cAMP.

The effect of a OP-3 on parathyroid hormone-mediated cAMP production in rat osteoblasts in vitro may be demonstrated as follows.

Rat osteoblasts are prepared and cultured in a multiwell plate as described above. The cultured cells then are divided into three groups: (1) wells which receive, for example, 1.0, 10.0 and 40.0 ng OP-3/ml medium); (2) wells which receive for example, TGF-β, at similar concentration ranges; and (3) a control group which receives no growth factors. The plate is then incubated for another 72 hours. At the end of the 72 hours the cells are treated with medium containing 0.5% bovine serum albumin (BSA) and 1mM 3-isobutyl-1-methylxanthine for 20 minutes followed by the addition into half of the wells of human recombinant parathyroid hormone (hPTH, Sigma, St. Louis) at a concentration of 200 ng/ml for 10 minutes. The cell layer then is extracted from each well with 0.5 ml of 1% Triton X-100. The cAMP levels then are determined using a radioimmunoassay kit (e.g., Amersham, Arlington Heights, Illinois). OP-3 alone stimulates an increase in the PTH-mediated cAMP response, and thus promotes the growth and expression of the osteoblast differentiated phenotype.

3.4 OP-3 Protein Induction of Osteocalcin Production

Osteocalcin is a bone-specific protein synthesized by osteoblasts which plays an integral role in the rate of bone mineralization in vivo. Circulating levels of osteocalcin in serum are used as a marker for osteoblast activity and bone formation in vivo. Induction of osteocalcin synthesis in osteoblast-enriched cultures can be used to demonstrate OP-3 morphogenic efficacy in vitro.

Rat osteoblasts are prepared and cultured in a multi-well plate as above. In this experiment the medium is supplemented with 10%FBS, and on day 2, cells are fed with fresh medium supplemented with fresh 10 mM β-glycerophosphate (Sigma, Inc.). Beginning on day 5 and twice weekly thereafter, cells are fed with a complete mineralization medium containing all of the above components plus fresh L(+)-ascorbate, at a final concentration of 50μg/ml medium. OP-3 then is added to the wells directly, e.g., in 50% acetonitrile (or 50% ethanol) containing 0.1% trifluoroacetic acid (TFA), at no more than 5μl morphogen/ml medium. Control wells receive solvent vehicle only. The cells then are re-fed and the conditioned medium sample diluted 1:1 in standard radioimmunoassay buffer containing standard protease inhibitors and stored at -20° C until assayed for osteocalcin. Osteocalcin synthesis is measured by standard radioimmunoassay using a commercially available osteocalcin-specific antibody.

Mineralization is determined on long term cultures (13 day) using a modified von Kossa staining technique on fixed cell layers: cells are fixed in fresh 4% paraformaldehyde at 23° C for 10 min, following rinsing cold 0.9% NaCl. Fixed cells then are stained for endogenous alkaline phosphatase at pH 9.5 for 10 min, using a commercially available kit (Sigma, Inc.) Purple stained cells then are dehydrated with methanol and air dried. after 30 min incubation in 3% $AgNO_3$ in the dark, H_2O-rinsed samples are exposed for 30 sec to 254 nm UV light to develop the black silver-stained phosphate nodules. Individual mineralized foci (at least 20 μm in size) are counted under a dissecting microscope and expressed as nodules/culture.

OP-3 stimulates osteocalcin synthesis in osteoblast cultures. The increased osteocalcin synthesis in response to OP-3 is dose dependent and shows a significant increase over the basal level after 13 days of incubation. The enhanced osteocalcin synthesis also can be confirmed by detecting the elevated osteocalcin mRNA message (20-fold increase) using a rat osteocalcin-specific probe. In addition, the increase in osteoclacin synthesis correlates with increased mineralization in long term osteoblast cultures as determined by the appearance of mineral nodules. OP-3 increases the initial mineralization rate significantly compared to untreated cultures.

3.5 Morphogen-Induced CAM Expression

The morphogens described herein induce CAM expression, particularly N-CAM expression, as part of their induction of morphogenesis (see copending USSN 922,813). CAMs are morphoregulatory molecules identified in all tissues as an essential step in tissue development. N-CAMs, which comprise at least 3 isoforms (N-CAM-180, N-CAM-140 and N-CAM-120, where "180", "140" and "120" indicate the apparent molecular weights of the isoforms as measured by SDS polyacrylamide gel electrophoresis) are expressed at least transiently in developing tissues, and permanently in nerve tissue. Both the N-CAM-180 and N-CAM-140 isoforms are expressed in both developing and adult tissue. The N-CAM-120 isoform is found only in adult tissue. Another neural CAM is L1.

The ability of OP-3 to stimulate CAM expression can be demonstrated using the following protocol, using NG108-15 cells. NG108-15 is a transformed hybrid cell line (neuroblastoma x glioma, America Type Tissue Culture (ATCC), Rockville, MD), exhibiting a morphology characteristic of transformed embryonic neurons. As described in Example 4, below, untreated NG108-15 cells exhibit a fibroblastic, or minimally differentiated, morphology and express only the 180 and 140 isoforms of N-CAM normally associated with a developing cell. Following morphogen treatment these cells exhibit a morphology characteristic of adult neurons and express enhanced levels of all three N-CAM isoforms.

In this example, NG108-15 cells are cultured for 4 days in the presence of increasing concentrations of OP-3 using standard culturing procedures, and standard Western blots performed on whole cell extracts. N-CAM isoforms are detected with an antibody which crossreacts with all three isoforms, mAb H28.123, obtained from Sigma Chemical Co., St. Louis, the different isoforms being distinguishable by their different mobilities on an electrophoresis gel. Control NG108-15 cells (untreated) express both the 140 kDa and the 180 kDa isoforms, but not the 120 kDa, as determined by Western blot analyses using up to 100 μg of protein. Treatment of NG108-15 cells with OP-3 results in a dose-dependent increase in the expression of the 180 kDa and 140 kDa isoforms, as well as the induction of the 120 kDa isoform induced. In addition, OP-3-induced CAM expression correlates with cell aggregation, as determined by histology.

Example 4. <u>OP-3 Protein-Induced Redifferentiation of Transformed Phenotype</u>

The OP-3 morphogens described herein also can induce redifferentiation of transformed cells to a morphology characteristic of untransformed cells. The examples provided below detail morphogen-induced redifferentiation of a transformed human cell line of neuronal origin (NG108-15); as well as mouse neuroblastoma cells (N1E-115), and human embryo carcinoma cells, to a morphology characteristic of untransformed cells.

As described above, NG108-15 is a transformed hybrid cell line produced by fusing neuroblastoma x glioma cells (obtained from ATTC, Rockville, MD), and exhibiting a morphology characteristic of transformed embryonic neurons, e.g., having a fibroblastic morphology. Specifically, the cells have polygonal cell bodies, short, spike-like processes and make few contacts with neighboring cells (see copending USSN 922,813). Incubation of NG108-15 cells, cultured in a chemically defined, serum-free medium, with 0.1 to 300 ng/ml of morphogen (e.g; OP-3) for four hours is anticipated to induce an orderly, dose-dependent change in cell morphology.

In the example, NG108-15 cells are subcultured on poly-L-lysine coated 6 well plates. Each well contains 40-50,000 cells in 2.5 ml of chemically defined medium. On the third day, 2.5 μl of morphogen (e.g., OP-3) in 60% ethanol containing 0.025% trifluoroacetic is added to each well. Morphogenic OP-3 of varying concentrations are tested (typically, concentration ranges of 0-300 ng/ml are tested). The media is changed daily with new aliquots of morphogen. OP-3 is anticipated to induce a dose-dependent redifferentiation of the transformed cells, including a rounding of the soma, an increase in phase brightness, extension of the short neurite processes, and other significant changes in the cellular ultrastructure. After several days treated cells should begin to form epithelioid sheets that then become highly packed, multi-layered aggregates, as determined visually by microscopic examination.

Moreover, morphogen-induced redifferentiation occurs without any associated changes in DNA synthesis, cell division, or cell viability, making it unlikely that the morphologic changes are secondary to cell differentiation or a toxic effect of the morphogen. In addition, the morphogen-induced redifferentiation does not inhibit cell division, as determined by ^{3}H-thymidine uptake, unlike other molecules which have been shown to stimulate differentiation of transformed cells, such as butyrate, DMSO, retanoic acid or Forskolin in analogous experiments. Thus, OP-3 maintains cell stability and viability after inducing redifferentiation.

The OP-3 morphogens described herein accordingly provide useful therapeutic agents for the treatment of neoplasias and neoplastic lesions of the nervous system, particularly in the treatment of neuroblastomas, including retinoblastomas, and gliomas.

As yet another, related example, the ability of OP-3 to induce the "redifferentiation" of transformed human cells may be demonstrated using the following assay. Specifically, the effect of OP-3 on human EC cells (embryo carcinoma cells, e.g., NTERA-Z CL.D1, ATCC, Rockville, MD) may be determined. In the absence of an external stimulant, these cells can be maintained as undifferentiated stem cells, and can be induced to grow in serum free media (SFM). In the absence of treatment with a morphogen, the cells proliferate

rampantly and are anchorage-independent. In the presence of morphogen, EC cells grow as flattened cells, becoming anchorage dependent and forming aggregates. In addition, growth rate is reduced approximately 10 fold. Ultimately, the cells are induced to differentiate. In the example, varying concentrations of OP-3 (e.g., 0-300 ng/ml) are added daily to cultured cells (e.g., 40-50,000 cells in 2.5 ml chemically defined medium), and the effects of treatment determined by visual examination. OP-3 is anticipated to stimulate redifferentiation of these cells to a morphology characteristic of untransformed embryo cells.

Example 5. Maintenance of Phenotype

Morphogenically active fragments of OP-3 also may be used to maintain a cell's differentiated phenotype. This application is particularly useful for inducing the continued expression of phenotype in senescent or quiescent cells.

5.1 In Vitro Model for Phenotypic Maintenance

The phenotypic maintenance capability of morphogens is determined readily. A number of differentiated cells become senescent or quiescent after multiple passages in vitro under standard tissue culture conditions well described in the art (e.g., Culture of Animal Cells: A Manual of Basic Techniques, C.R.

Freshney, ed., Wiley, 1987). However, if these cells are cultivated in vitro in association with a morphogen such as OP-3, cells are stimulated to maintain expression of their phenotype through multiple passages. For example, the alkaline phosphatase activity of cultured osteoblasts, such as cultured osteosarcoma cells and calvaria cells, is significantly reduced after multiple passages in vitro. However, if the cells are cultivated in the presence of OP-3, alkaline phosphatase activity should be maintained over extended periods of time. Similarly, phenotypic expression of myocytes also is maintained in the presence of a morphogen. In the experiment, osteoblasts are cultured as described in Example 2. The cells are divided into groups, incubated with varying concentrations of OP-3 (e.g., 0-300 ng/ml) and passaged multiple times (e.g., 3-5 times) using standard methodology. Passaged cells then are tested for alkaline phosphatase activity, as described in Example 3 as an indication of differentiated cell metabolic function. Osteoblasts cultured in the absence of OP-3 should have reduced alkaline phosphatase activity, as compared to OP-3-treated cells.

5.2 In Vivo Model for Phenotypic Maintenance

Phenotypic maintenance capability also may be demonstrated in vivo, using a rat model for osteoporosis, as disclosed in international application PCT/US92/07432 (WO93/05751). As described therein,

Long Evans female rats (Charles River Laboratories, Wilmington, MA) are Sham-operated (control animals) or ovariectomized using standard surgical techniques, to produce an osteoporotic condition resulting from decreased estrogen production. Shortly following surgery, e.g., 200 days after ovariectomy, rats are systemically provided with phosphate buffered saline (PBS) or morphogen, (e.g., OP-3, 1-100 μg) for 21 days (e.g., by daily tail vein injection.) The rats then are sacrificed and serum alkaline phosphatase levels, serum calcium levels, and serum osteocalcin levels are determined, using standard methodologies as described therein and above. Elevated levels of osteocalcin and alkaline phosphatase should be observed in the rats treated with an effective amount of OP-3. Moreover, histomorphometric analysis on the tibial diasypheal bone is anticipated to show improved bone mass in OP-3-treated animals as compared with untreated, ovariectomized rats. In fact, the bone mass of OP-3-animals is anticipated to be comparable to (e.g., approaches) that of the sham-operated (e.g., nonovarectomized) rats.

Example 6. Proliferation of Progenitor Cell Populations

Progenitor cells may be stimulated to proliferate in vivo or ex vivo. The cells may be stimulated in vivo by injecting or otherwise providing a sterile preparation containing the morphogenically active fragment of OP-3 into the individual. For example, the

hemopoietic pluripotential stem cell population of an individual may be stimulated to proliferate by injecting or otherwise providing an appropriate concentration of OP-3 to the individual's bone marrow.

Progenitor cells may be stimulated ex vivo by contacting progenitor cells of the population to be enhanced with a morphogenically active fragment of OP-3 under sterile conditions at a concentration and for a time sufficient to stimulate proliferation of the cells. Suitable concentrations and stimulation times may be determined empirically, essentially following the procedure described in Example 2, above. A morphogen concentration of between about 0.1-100 ng/ml and a stimulation period of from about 10 minutes to about 72 hours, or, more generally, about 24 hours, typically should be sufficient to stimulate a cell population of about 10^4 to 10^6 cells. The stimulated cells then are provided to the individual as, for example, by injecting the cells to an appropriate in vivo locus. Suitable biocompatible progenitor cells may be obtained by any of the methods known in the art or described hereinabove.

Example 7. Regeneration of Damaged or Diseased Tissue

OP-3 may be used to repair diseased or damaged mammalian tissue. The tissue to be repaired preferably is assessed first, and excess necrotic or interfering scar tissue removed as needed, e.g., by ablation or by surgical, chemical, or other methods known in the medical arts.

OP-3 then may be provided directly to the tissue locus as part of a sterile, biocompatible composition, either by surgical implantation or injection. The morphogen also may be provided systemically, as by oral or parenteral administration. Alternatively, a sterile, biocompatible composition containing progenitor cells stimulated by a morphogenically active fragment of OP-3 may be provided to the tissue locus. The existing tissue at the locus, whether diseased or damaged, provides the appropriate matrix to allow the proliferation and tissue-specific differentiation of progenitor cells. In addition, a damaged or diseased tissue locus, particularly one that has been further assaulted by surgical means, provides a morphogenically permissive environment. Systemic provision of OP-3 should be sufficient for certain applications (e.g., in the treatment of osteoporosis and other disorders of the bone remodeling cycle, as an example).

In some circumstances, particularly where tissue damage is extensive, the tissue may not be capable of providing a sufficient matrix for cell influx and proliferation. In these instances, it may be necessary to provide OP-3 or progenitor cells stimulated by OP-3 to the tissue locus in association with a suitable, biocompatible, formulated matrix, prepared by any of the means described below. The matrix preferably is <u>in vivo</u> biodegradable. The matrix also may be tissue-specific and/or may comprise porous particles having dimensions within the range of 70-850μm, most preferably 150-420μm.

OP-3 also may be used to prevent or substantially inhibit immune/inflammatory response-mediated tissue damage and scar tissue formation following an injury. OP-3 is provided to a newly injured tissue locus, to induce tissue morphogenesis at the locus, preventing the aggregation of migrating fibroblasts into non-differentiated connective tissue. OP-3 preferably is provided as a sterile pharmaceutical preparation injected into the tissue locus within five hours of the injury. Where an immune/inflammatory response is unavoidably or deliberately induced, as part of, for example, a surgical or other aggressive clinical therapy, OP-3 preferably is provided prophylactically to the patient, prior to, or concomitant with, the therapy.

Below are several examples, describing protocols for demonstrating OP-3-induced tissue morphogenesis in bone, liver, nerve, dentin, cementum and periodontal tissue.

7.1 OP-3-Induced Bone Morphogenesis

A particularly useful mammalian tissue model system for demonstrating and evaluating the morphogenic activity of a protein is the endochondral bone tissue morphogenesis model known in the art and described, for example, in U.S. Pat. No. 4,968,590. The ability to induce endochondral bone formation includes the ability to induce the proliferation of progenitor cells into

chondroblasts and osteoblasts, the ability to induce cartilage matrix formation, cartilage calcification, and bone remodeling, and the ability to induce formation of an appropriate vascular supply and hematopoeitic bone marrow differentiation.

The local environment in which the morphogenic material is placed is important for tissue morphogenesis. As used herein, "local environment" is understood to include the tissue structural matrix and the environment surrounding the tissue. For example, in addition to needing an appropriate anchoring substratum for their proliferation, the cells stimulated by morphogens need signals to direct the tissue-specificity of their differentiation. These signals vary for the different tissues and may include cell surface markers. In addition, vascularization of new tissue requires a local environment which supports vascularization.

The following sets forth various procedures for evaluating the in vivo morphogenic utility of OP-3 and OP-3-containing compositions. The compositions may be injected or surgically implanted in a mammal, following any of a number of procedures well known in the art. For example, surgical implant bioassays may be performed essentially following the procedure of Sampath et al. (1983) PNAS 80:6591-6595 and U.S. Pat No. 4,968,590.

Histological sectioning and staining is preferred to determine the extent of morphogenesis in vivo, particularly in tissue repair procedures. Excised implants are fixed in Bouins Solution, embedded in paraffin, and cut into 6-8 μm sections. Staining with toluidine blue or hemotoxylin/eosin demonstrates clearly the ultimate development of the new tissue. Twelve day implants are usually sufficient to determine whether the implants contain newly induced tissue.

Successful implants exhibit a controlled progression through the stages of induced tissue development allowing one to identify and follow the tissue-specific events that occur. For example, in endochondral bone formation the stages include: (1) leukocytes on day one; (2) mesenchymal cell migration and proliferation on days two and three; (3) chondrocyte appearance on days five and six; (4) cartilage matrix formation on day seven; (5) cartilage calcification on day eight; (6) vascular invasion, appearance of osteoblasts, and formation of new bone on days nine and ten; (7) appearance of osteoclastic cells, and the commencement of bone remodeling and dissolution of the implanted matrix on days twelve to eighteen; and (8) hematopoietic bone marrow differentiation in the resulting ossicles on day twenty-one.

In addition to histological evaluation, biological markers may be used as markers for tissue morphogenesis. Useful markers include tissue-specific

enzymes whose activities may be assayed (e.g., spectrophotometrically) after homogenization of the implant. These assays may be useful for quantitation and for rapidly obtaining an estimate of tissue formation after the implants are removed from the animal. For example, alkaline phosphatase activity may be used as a marker for osteogenesis.

Incorporation of systemically provided OP-3 may be followed using tagged fragments (e.g., radioactively labelled) and determining their localization in the new tissue, and/or by monitoring their disappearance from the circulatory system using a standard labeling protocol and pulse-chase procedure. OP-3 also may be provided with a tissue-specific molecular tag, whose uptake may be monitored and correlated with the concentration of OP-3 provided. As an example, ovary removal in female rats results in reduced bone alkaline phosphatase activity, and renders the rats predisposed to osteoporosis (as described in Example 5). If the female rats now are provided with OP-3, a reduction in the systemic concentration of calcium should be seen, which correlates with the presence of the provided OP-3 and which is anticipated to correspond with increased alkaline phosphatase activity.

7.2 Morphogen-Induced Liver Regeneration

As another example, a method for inducing morphogenesis of substantially injured liver tissue following a partial hepatectomy utilizing OP-3 is

presented. Variations on this general protocol may be used to test morphogen activity of OP-3 in other different tissues. The general method involves excising an essentially nonregenerating portion of a tissue and providing OP-3, preferably as a soluble pharmaceutical preparation to the excised tissue locus, closing the wound, and examining the site at a future date. Like bone, liver has a potential to regenerate upon injury during post-fetal life.

OP-3, e.g., 1 mg/ml, in a biocompatible solution, for example, (e.g., a purified recombinant mature form of OP-3, is solubilized in 50% ethanol, or compatible solvent, containing 0.1% trifluoroacetic acid, or compatible acid. Alternatively, the mature protein may be solubilized by association with a pro domain. The injectable OP-3 solution is prepared, e.g., by diluting one volume of OP-3 solvent-acid stock solution with 9 volumes of 0.2% rat serum albumin in sterile PBS (phosphate-buffered saline).

In the experiment, growing rats or aged rats (e.g., Long Evans, Charles River Laboratories, Wilmington) are anesthetized by using ketamine. Two of the liver lobes (left and right) are cut out (approximately 1/3 of the lobe) and the OP-3 is injected locally at multiple sites along the cut ends. The amount of OP-3 injected may be, e.g., 100 μg in 1000 μl of PBS/RSA (phosphate buffered saline/rat serum albumin) injection buffer. Placebo samples are injection buffer only. In experimental essays, five rats in each group preferably are used. The wound is closed and the rats are allowed to eat normal food and drink tap water.

After 12 days, the rats are sacrificed and liver regeneration is observed visually, to evaluate the effects of the OP-3 on liver regeneration most effectively. The OP-3 fragment-injected group is anticipated to show, e.g., complete liver tissue regeneration with no sign remaining of any cut in the liver. By contrast, the control group into which only PBS is injected, show only minimal regeneration with the incision remaining in the sample. Previous experiments with other morphogens (e.g., OP-1) show these morphogens alone induce liver tissue regeneration.

7.3 Morphogen-Induced Dentin, Cementum and Periodontal Ligament Regeneration

As still another example, the ability of OP-3 to induce dentinogenesis also may be demonstrated. To date, the unpredictable response of dental pulp tissue to injury is a basic clinical problem in dentistry. Cynomolgus monkeys are chosen as primate models as monkeys are presumed to be more indicative of human dental biology than models based on lower non-primate mammals.

Using standard dental surgical procedures, small areas (e.g., 2mm) of dental pulps are surgically exposed by removing the enamel and dentin immediately above the pulp (by drilling) of sample teeth, performing a partial amputation of the coronal pulp tissue, inducing hemostasis, application of the pulp treatment, and sealing and filling the cavity by standard procedures.

Pulp treatments used may include: a morphogenically active fragment of OP-3 dispersed in a carrier matrix; carrier matrix alone, and no treatment. Twelve teeth per animal (four for each treatment) are prepared, and two animals are used. At four weeks, teeth are extracted and processed histologically for analysis of dentin formation, and/or ground to analyze dentin mineralization. The effect of OP-3 on osteodentin reparation may be observed visually by comparing control samples treatment (PBS) with OP-3. OP-3 plus a carrier matrix induces formation of reparative or osteodentin bridges on surgically exposed healthy dental pulps. By contrast, pulps treated with carrier matrix alone, do not form reparative dentin.

Similarly, implanting demineralized teeth and OP-3 into surgically prepared canine tooth sockets is anticipated to stimulate new periodontal tissue formation, including new cementum and periodontal ligament, as well as new alveolar bone and dentin tissue, as described for OP-1 in international application PCT/US92/08742, filed 9/15/93. By contrast, untreated teeth or teeth treated with carrier vehicle alone do not induce periodontal tissue growth.

7.4 Morphogen-Induced Nerve Tissue Repair

As yet another example, the induction of regenerative effects on central nervous system (CNS) repair, by a morphogenically active fragment of OP-3, may be demonstrated using a rat brain stab model. In the experiment, male Long Evans rats are anesthetized

and the head area prepared for surgery. The calvariae is exposed using standard surgical procedures and a hole drilled toward the center of each lobe using a 0.035K wire, just piercing the calvariae. 25µl solutions containing either morphogen (e.g., OP-3, 25µg) or PBS then is provided to each of the holes by Hamilton syringe. Solutions are delivered to a depth approximately 3 mm below the surface, into the underlying cortex, corpus callosum and hippocampus. The skin then is sutured and the animal allowed to recover.

Three days post surgery, rats are sacrificed by decapitation and their brains processed for sectioning. Scar tissue formation is evaluated by immunofluoresence staining for glial fibrillary acidic protein, a marker protein for glial scarring, to qualitatively determine the degree of scar formation. Sections also are probed with OP-3-specific antibody to determine the presence of the protein. Reduced levels of glial fibrillary acidic protein are anticipated to be observed in the tissue sections of animals treated with OP-3, evidencing the ability of the morphogen to inhibit glial scar formation, thereby stimulating nerve regeneration.

The ability of OP-3 to stimulate peripheral nervous system axonal growth over extended distances may be demonstrated using the following model. Neurons of the peripheral nervous system can sprout new processes on their own following injury, but without guidance these sproutings typically fail to connect appropriately and

die. Where the break is extensive, e.g., greater than 5 or 10 mm, regeneration is poor or nonexistent. Previous experiments with other morphogens, e.g., OP-1, show that morphogens stimulate peripheral nervous system axonal growth over extended distances, allowing repair and regeneration of damaged peripheral neural pathways.

In this example OP-3 stimulation of nerve regeneration is demonstrated using the rat sciatic nerve model. The rat sciatic nerve can regenerate spontaneously across a 5 mm gap, and occasionally across a 10 mm gap, provided that the severed ends are inserted in a saline-filled nerve guidance channel. In this experiment, nerve regeneration across at least a 12mm gap is tested.

Adult female Sprague-Dawley rats (Charles River Laboratories, Inc.) weighing 230-250 g are anesthetized with intraperitoneal injections of sodium pentobarbital (35 mg/kg body weight). A skin incision is made parallel and just posterior to the femur. The avascular intermuscular plane between vastus lateralis and hamstring muscles are entered and followed to the loose fibroareolar tissue surrounding the sciatic nerve. The loose tissue is divided longitudinally thereby freeing the sciatic nerve over its full extent without devascularizing any portion. Under a surgical microscope the sciatic nerves are transected with

microscissors at mid-thigh and grafted with a OP-3 gel graft that separates the nerve stumps by 12 mm. The graft region is encased in a silicone tube 20 mm in length with a 1.5 mm inner diameter, the interior of which is filled with the morphogen solution. Specifically, the central 12 mm of the tube consists of an OP-3 gel prepared by mixing 1 to 5 μg of substantially pure recombinantly produced OP-3 protein with approximately 100 μl of MATRIGEL™ (from Collaborative Research, Inc., Bedford, MA), an extracellular matrix extract derived from mouse sarcoma tissue, and containing solubilized tissue basement membrane, including laminin, type IV collagen, heparin sulfate, proteoglycan and entactin, in phosphate-buffered saline. The morphogen-filled tube then is implanted directly into the defect site, allowing 4 mm on each end to insert the nerve stumps. Each stump is abutted against the morphogen gel and is secured in the silicone tube by three stitches of commercially available surgical 10-0 nylon through the epineurium, the fascicle protective sheath.

In addition to OP-3 gel grafts, control grafts of empty silicone tubes, silicone tubes filled with gel only and "reverse" autografts, wherein 12 mm transected segments of the animal's sciatic nerve are rotated 180° prior to suturing, preferably also are grafted. All experiments preferably are performed in quadruplicate. All wounds preferably are closed by wound clips that are removed after 10 days. Rats can be grafted on both

legs. At 3 weeks the animals are sacrificed, and the grafted segments removed and frozen on dry ice immediately. Frozen sections then are cut throughout the graft site, and examined for axonal regeneration by immunofluorescent staining using anti-neurofilament antibodies labeled with flurocein (obtained, for example, from Sigma Chemical Co., St. Louis).

Regeneration of the sciatic nerve is anticipated to occur across the entire 12 mm distance in all graft sites wherein the gap is filled with the OP-3 gel. By contrast, empty silicone tubes, gel alone and reverse autografts do not show nerve regeneration.

Example 8. Identification of Morphogen-Expressing Tissue

Determining the tissue distribution of morphogens may be used to identify different morphogens expressed in a given tissue, as well as to identify new, related morphogens. Tissue distribution also may be used to identify useful morphogen-producing tissue for use in screening and identifying candidate morphogen-stimulating agents. The morphogens (or their mRNA transcripts) readily are identified in different tissues using standard methodologies and minor modifications thereof in tissues where expression may be low. For example, protein distribution may be determined using standard Western blot analysis or immunofluorescent techniques, and antibodies specific

to the morphogen or morphogens of interest. Similarly, the distribution of morphogen transcripts may be determined using standard Northern hybridization protocols and transcript-specific probes.

Any probe capable of hybridizing specifically to a transcript, and distinguishing the transcript of interest from other, related transcripts may be used. Because the morphogens described herein share such high sequence homology in their active, C-terminal domains, the tissue distribution of a specific morphogen transcript may best be determined using a probe specific for the pro region of the immature protein and/or the N-terminal region of the mature protein. Another useful sequence is the 3' non-coding region flanking and immediately following the stop codon. These portions of the sequence vary substantially among the morphogens of this invention, and accordingly, are specific for each protein. For example, a particularly useful OP-3-specific probe sequence is one derived from a portion of the 3' untranslated sequence, e.g., nucleotides 1310-1674 of Seq. ID No. 1, which shares little or no homology with other morphogen sequences, including OP-2. The chosen fragment then is labelled using standard means well known and described in the art.

Using these morphogen-specific probes, which may be synthetically engineered or obtained from cloned sequences, morphogen transcripts can be identified in mammalian tissue, using standard methodologies well

known to those having ordinary skill in the art. A detailed description of a suitable hybridization protocol is described in Ozkaynak, et al., (1991) Biochem. Biophys. Res. Commn. 179:116-123, and Ozkaynak, et al. (1992) J. Biol. Chemistry 267:25220-25227. Briefly, total RNA is prepared from various tissues (e.g., murine embryo and developing and adult liver, kidney, testis, heart, brain, thymus, stomach) by a standard methodology such as by the method of Chomczyaski et al. ((1987) Anal. Biochem 162:156-159) and described below. Poly (A)+ RNA is prepared by using oligo (dT)-cellulose chromatography (e.g., Type 7, from Pharmacia LKB Biotechnology, Inc.). Poly (A)+ RNA (generally 15 μg) from each tissue is fractionated on a 1% agarose/formaldehyde gel and transferred onto a Nytran membrane (Schleicher & Schuell). Following the transfer, the membrane is baked at 80°C and the RNA is cross-linked under UV light (generally 30 seconds at 1 mW/cm^2). Prior to hybridization, the appropriate probe is denatured by heating. The hybridization is carried out in a lucite cylinder rotating in a roller bottle apparatus at approximately 1 rev/min for approximately 15 hours at 37°C using a hybridization mix of 40% formamide, 5 x Denhardts, 5 x SSPE, and 0.1% SDS. Following hybridization, the non-specific counts are washed off the filters in 0.1 x SSPE, 0.1% SDS at 50°C.

An OP-3-specific 0.5 kb probe was made from a StuI-BglII fragment of OP-3 cDNA. The fragment contains the 3' untranslated sequence from nucleotides 1310-1674, plus an additional 140 bases. The fragment was

labelled using standard techniques and the hybridization performed as described. To date, OP-3, like OP-2, appears to be expressed primarily in early embryonic tissue. Specifically, Northern blots of murine embryos show abundant OP-3 expression in 8-day embryos, demonstrated by a strong band at 2.9 kb and a weaker band at 2.3 kb.

Example 9. Screening Assay for Candidate Compounds which Alter Endogenous Morphogen Levels

Candidate compound(s) which may be administered to affect the level of endogenous OP-3 morphogen may be found using the following screening assay, in which the level of OP-3 production by a cell type which produces measurable levels of the morphogen is determined with and without incubating the cell in culture with the compound, in order to assess the effects of the compound on the cell. This can be accomplished by detection of the morphogen either at the protein or RNA level. A detailed description also may be found in international application PCT/US92/07359, (WO93/05172).

9.1 Growth of Cells in Culture

Cell cultures of kidney, adrenals, urinary bladder, brain, or other organs, may be prepared as described widely in the literature. For example, kidneys may be explanted from neonatal or new born or young or adult rodents (mouse or rat) and used in organ culture as whole or sliced (1-4 mm) tissues. Primary tissue

cultures and established cell lines, also derived from kidney, adrenals, urinary, bladder, brain, mammary, or other tissues may be established in multiwell plates (6 well or 24 well) according to conventional cell culture techniques, and are cultured in the absence or presence of serum for a period of time (1-7 days). Cells may be cultured, for example, in Dulbecco's Modified Eagle medium (Gibco, Long Island, NY) containing serum (e.g., fetal calf serum at 1%-10%, Gibco) or in serum-deprived medium, as desired, or in defined medium (e.g., containing insulin, transferrin, glucose, albumin, or other growth factors).

Samples for testing the level of morphogen production includes culture supernatants or cell lysates, collected periodically and evaluated for morphogen production by immunoblot analysis (Sambrook et al., eds., 1989, Molecular Cloning, Cold Spring Harbor Press, Cold Spring Harbor, NY), or a portion of the cell culture itself, collected periodically and used to prepare polyA+ RNA for RNA analysis. To monitor de novo morphogen synthesis, some cultures are labeled according to conventional procedures with an ^{35}S-methionine/^{35}S-cysteine mixture for 6-24 hours and then evaluated for morphogenic protein synthesis by conventional immunoprecipitation methods.

9.2 Determination of Level of Morphogenic Protein

In order to quantitate the production of a morphogenic protein, e.g., OP-3, by a cell type, an immunoassay may be performed to detect the morphogen

using a polyclonal or monoclonal antibody specific for that protein. For example, OP-3 may be detected using a polyclonal antibody specific for OP-3 in an ELISA, as follows.

1 μg/100 μl of affinity-purified polyclonal rabbit IgG specific for OP-3 is added to each well of a 96-well plate and incubated at 37°C for an hour. The wells are washed four times with 0.167M sodium borate buffer with 0.15 M NaCl (BSB), pH 8.2, containing 0.1% Tween 20. To minimize non-specific binding, the wells are blocked by filling completely with 1% bovine serum albumin (BSA) in BSB and incubating for 1 hour at 37°C. The wells are then washed four times with BSB containing 0.1% Tween 20. A 100 μl aliquot of an appropriate dilution of each of the test samples of cell culture supernatant is added to each well in triplicate and incubated at 37°C for 30 min. After incubation, 100 μl biotinylated rabbit anti-OP-3 serum (stock solution is about 1 mg/ml and diluted 1:400 in BSB containing 1% BSA before use) is added to each well and incubated at 37°C for 30 min. The wells are then washed four times with BSB containing 0.1% Tween 20. 100 μl strepavidin-alkaline (Southern Biotechnology Associates, Inc. Birmingham, Alabama, diluted 1:2000 in BSB containing 0.1% Tween 20 before use) is added to each well and incubated at 37°C for 30 min. The plates are washed four times with 0.5M Tris buffered Saline (TBS), pH 7.2. 50μl substrate (ELISA Amplification System Kit, Life Technologies, Inc., Bethesda, MD) is added to each well incubated at room temperature for 15

min. Then, 50 µl amplifier (from the same amplification system kit) is added and incubated for another 15 min at room temperature. The reaction is stopped by the addition of 50 µl 0.3 M sulphuric acid. The OD at 490 nm of the solution in each well is recorded. To quantitate OP-3 in culture media, an OP-3 standard curve is performed in parallel with the test samples.

Polyclonal antibody may be prepared as follows. Each rabbit is given a primary immunization of 100 ug/500 µl recombinantly-produced OP-3 protein or protein fragment in 0.1% SDS mixed with 500 µl Complete Freund's Adjuvant. The antigen is injected subcutaneously at multiple sites on the back and flanks of the animal. The rabbit is boosted after a month in the same manner using incomplete Freund's Adjuvant. Test bleeds are taken from the ear vein seven days later. Two additional boosts and test bleeds are performed at monthly intervals until antibody against OP-3 is detected in the serum using an ELISA assay. Then, the rabbit is boosted monthly with 100 µg of antigen and bled (15 ml per bleed) at days seven and ten after boosting.

Monoclonal antibody specific for a given morphogen may be prepared as follows. A mouse is given two injections of OP-3 protein or a protein fragment specific for OP-3. The protein preferably is recombinantly produced. The first injection contains 100µg of OP-3 in complete Freund's adjuvant and is given subcutaneously. The second injection contains 50 µg of OP-3 in incomplete adjuvant and is given intraperitoneally. The mouse then receives a total of

230 μg of OP-3 in four intraperitoneal injections at various times over an eight month period. One week prior to fusion, the mouse is boosted intraperitoneally with OP-3 (e.g., 100 μg) and may be additionally boosted with an OP-3-specific peptide (e.g., corresponding to the N-terminus of the mature protein) conjugated to bovine serum albumin with a suitable crosslinking agent. This boost can be repeated five days (IP), four days (IP), three days (IP) and one day (IV) prior to fusion. The mouse spleen cells then are fused to commercially available myeloma cells at a ratio of 1:1 using PEG 1500 (Boeringer Mannheim, Germany), and the fused cells plated and screened for OP-3-specific antibodies using OP-3 as antigen. The cell fusion and monoclonal screening steps readily are performed according to standard procedures well described in standard texts widely available in the art.

Other Embodiments

The invention may be embodied in other specific forms without departing from the spirit or essential characteristics thereof. The present embodiments are therefore to be considered in all respects as illustrative and not restrictive, the scope of the invention being indicated by the appended claims rather than by the foregoing description, and all changes which come within the meaning and range of equivalency of the claims are therefore intended to be embraced therein.

SEQUENCE LISTING

(1) GENERAL INFORMATION:

(i) APPLICANT:
(A) NAME: CREATIVE BIOMOLECULES, INC.
(B) STREET: 45 SOUTH STREET
(C) CITY: HOPKINTON
(D) STATE: MA
(E) COUNTRY: USA
(F) POSTAL CODE (ZIP): 01748
(G) TELEPHONE: 1-508-435-9001
(H) TELEFAX: 1-508-435-0454
(I) TELEX:

(ii) TITLE OF INVENTION: OP3-INDUCED MORPHOGENESIS

(iii) NUMBER OF SEQUENCES: 13

(iv) CORRESPONDENCE ADDRESS:
(A) ADDRESSEE: CREATIVE BIOMOLECULES, INC.
(B) STREET: 45 SOUTH STREET
(C) CITY: HOPKINTON
(D) STATE: MA
(E) COUNTRY: USA
(F) ZIP: 01748

(v) COMPUTER READABLE FORM:
(A) MEDIUM TYPE: Floppy disk
(B) COMPUTER: IBM PC compatible
(C) OPERATING SYSTEM: PC-DOS/MS-DOS
(D) SOFTWARE: PatentIn Release #1.0, Version #1.25

(vi) CURRENT APPLICATION DATA:
(A) APPLICATION NUMBER:
(B) FILING DATE:
(C) CLASSIFICATION:

(vii) PRIOR APPLICATION DATA:
(A) APPLICATION NUMBER: US 07/667,274
(B) FILING DATE: 11-MAR-1991

(vii) PRIOR APPLICATION DATA:
(A) APPLICATION NUMBER: US 07/752,764
(B) FILING DATE: 30-AUG-1991

(vii) PRIOR APPLICATION DATA:
(A) APPLICATION NUMBER: US 07/753,059
(B) FILING DATE: 30-AUG-1991

(vii) PRIOR APPLICATION DATA:
(A) APPLICATION NUMBER: US 07/752,857
(B) FILING DATE: 30-AUG-1991

(vii) PRIOR APPLICATION DATA:
(A) APPLICATION NUMBER: US 07/923,780
(B) FILING DATE: 31-JUL-1992

(vii) PRIOR APPLICATION DATA:
(A) APPLICATION NUMBER: US 07/922,813
(B) FILING DATE: 31-JUL-1992

(viii) ATTORNEY/AGENT INFORMATION:
(A) NAME: PITCHER ESQ, EDMUND R
(B) REGISTRATION NUMBER: 27,829
(C) REFERENCE/DOCKET NUMBER: CRP-076PC

(ix) TELECOMMUNICATION INFORMATION:
(A) TELEPHONE: (508)435-9001

(2) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1674 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(ix) FEATURE:
(A) NAME/KEY: CDS
(B) LOCATION: 69..1268
(D) OTHER INFORMATION: /note= "mOP3-PP"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

GGATCCGCGG CGCTGTCCCA TCCTTGTCGT CGAGGCGTCG CTGGATGCGA GTCCGCTAAA 60

CGTCCGAG ATG GCT GCG CGT CCG GGA CTC CTA TGG CTA CTG GGC CTG GCT 110
Met Ala Ala Arg Pro Gly Leu Leu Trp Leu Leu Gly Leu Ala
1 5 10

CTG TGC GTG TTG GGC GGC GGT CAC CTC TCG CAT CCC CCG CAC GTC TTT 158
Leu Cys Val Leu Gly Gly Gly His Leu Ser His Pro Pro His Val Phe
15 20 25 30

```
CCC CAG CGT CGA CTA GGA GTA CGC GAG CCC CGC GAC ATG CAG CGC GAG                206
Pro Gln Arg Arg Leu Gly Val Arg Glu Pro Arg Asp Met Gln Arg Glu
                35                  40                  45

ATT CGG GAG GTG CTG GGG CTA GCC GGG CGG CCC CGA TCC CGA GCA CCG               254
Ile Arg Glu Val Leu Gly Leu Ala Gly Arg Pro Arg Ser Arg Ala Pro
            50                  55                  60

GTC GGG GCT GCC CAG CAG CCA GCG TCT GCG CCC CTC TTT ATG TTG GAC               302
Val Gly Ala Ala Gln Gln Pro Ala Ser Ala Pro Leu Phe Met Leu Asp
        65                  70                  75

CTG TAC CGT GCC ATG ACG GAT GAC AGT GGC GGT GGG ACC CCG CAG CCT               350
Leu Tyr Arg Ala Met Thr Asp Asp Ser Gly Gly Gly Thr Pro Gln Pro
    80                  85                  90

CAC TTG GAC CGT GCT GAC CTG ATT ATG AGC TTT GTC AAC ATA GTG GAA               398
His Leu Asp Arg Ala Asp Leu Ile Met Ser Phe Val Asn Ile Val Glu
 95                  100                 105                 110

CGC GAC CGT ACC CTG GGC TAC CAG GAG CCA CAC TGG AAG GAA TTC CAC               446
Arg Asp Arg Thr Leu Gly Tyr Gln Glu Pro His Trp Lys Glu Phe His
                115                 120                 125

TTT GAC CTA ACC CAG ATC CCT GCT GGG GAG GCT GTC ACA GCT GCT GAG               494
Phe Asp Leu Thr Gln Ile Pro Ala Gly Glu Ala Val Thr Ala Ala Glu
            130                 135                 140

TTC CGG ATC TAC AAA GAA CCC AGT ACC CAC CCG CTC AAC ACA ACC CTC               542
Phe Arg Ile Tyr Lys Glu Pro Ser Thr His Pro Leu Asn Thr Thr Leu
        145                 150                 155

CAC ATC AGC ATG TTC GAA GTG GTC CAA GAG CAC TCC AAC AGG GAG TCT               590
His Ile Ser Met Phe Glu Val Val Gln Glu His Ser Asn Arg Glu Ser
    160                 165                 170

GAC TTG TTC TTT TTG GAT CTT CAG ACG CTC CGA TCT GGG GAC GAG GGC               638
Asp Leu Phe Phe Leu Asp Leu Gln Thr Leu Arg Ser Gly Asp Glu Gly
175                 180                 185                 190

TGG CTG GTG CTG GAC ATC ACA GCA GCC AGT GAC CGA TGG CTG CTG AAC               686
Trp Leu Val Leu Asp Ile Thr Ala Ala Ser Asp Arg Trp Leu Leu Asn
                195                 200                 205

CAT CAC AAG GAC CTA GGA CTC CGC CTC TAT GTG GAA ACC GAG GAT GGG               734
His His Lys Asp Leu Gly Leu Arg Leu Tyr Val Glu Thr Glu Asp Gly
            210                 215                 220
```

```
CAC AGC ATA GAT CCT GGC CTA GCT GGT CTG CTT GGA CGA CAA GCA CCA      782
His Ser Ile Asp Pro Gly Leu Ala Gly Leu Leu Gly Arg Gln Ala Pro
        225                 230                 235

CGC TCC AGA CAG CCT TTC ATG GTT GGT TTC TTC AGG GCC AAC CAG AGT      830
Arg Ser Arg Gln Pro Phe Met Val Gly Phe Phe Arg Ala Asn Gln Ser
    240                 245                 250

CCT GTG CGG GCC CCT CGA ACA GCA AGA CCA CTG AAG AAG AAG CAG CTA      878
Pro Val Arg Ala Pro Arg Thr Ala Arg Pro Leu Lys Lys Lys Gln Leu
255                 260                 265                 270

AAT CAA ATC AAC CAG CTG CCG CAC TCC AAC AAA CAC CTA GGA ATC CTT      926
Asn Gln Ile Asn Gln Leu Pro His Ser Asn Lys His Leu Gly Ile Leu
                275                 280                 285

GAT GAT GGC CAC GGT TCT CAC GGC AGA GAA GTT TGC CGC AGG CAT GAG      974
Asp Asp Gly His Gly Ser His Gly Arg Glu Val Cys Arg Arg His Glu
            290                 295                 300

CTC TAT GTC AGC TTC CGT GAC CTT GGC TGG CTG GAC TCT GTC ATT GCC      1022
Leu Tyr Val Ser Phe Arg Asp Leu Gly Trp Leu Asp Ser Val Ile Ala
        305                 310                 315

CCC CAG GGC TAC TCC GCC TAT TAC TGT GCT GGG GAG TGC ATC TAC CCA      1070
Pro Gln Gly Tyr Ser Ala Tyr Tyr Cys Ala Gly Glu Cys Ile Tyr Pro
    320                 325                 330

CTG AAC TCC TGT ATG AAC TCC ACC AAC CAC GCC ACT ATG CAG GCC CTG      1118
Leu Asn Ser Cys Met Asn Ser Thr Asn His Ala Thr Met Gln Ala Leu
335                 340                 345                 350

GTA CAT CTG ATG AAG CCA GAT ATC ATC CCC AAG GTG TGC TGT GTG CCT      1166
Val His Leu Met Lys Pro Asp Ile Ile Pro Lys Val Cys Cys Val Pro
                355                 360                 365

ACT GAG CTG AGT GCC ATT TCT CTG CTC TAC TAT GAT AGA AAC AAT AAT      1214
Thr Glu Leu Ser Ala Ile Ser Leu Leu Tyr Tyr Asp Arg Asn Asn Asn
            370                 375                 380

GTC ATC CTG CGC AGG GAG CGC AAC ATG GTA GTC CAG GCC TGT GGC TGC      1262
Val Ile Leu Arg Arg Glu Arg Asn Met Val Val Gln Ala Cys Gly Cys
        385                 390                 395

CAC TGAGTCCCTG CCCAACAGCC TGCTGCCATC CCATCTATCT AGTCAGGCCT           1315
His
    400

CTCTTCCAAG GCAGGAAACC AACAAAGAGG GAAGGCAGTG CTTTCAACTC CATGTCCACA    1375
```

TTCACAGTCT TGGCCCTCTC TGTTCTTTTT GCCAAGGCTG AGAAGATGGT CCTAGTTATA 1435

ACCCTGGTGA CCTCAGTAGC CCGATCTCTC ATCTCCCCAA ACTCCCCAAT GCAGCCAGGG 1495

GCATCTATGT CCTTTGGGAT TGGGCACAGA AGTCCAATTT ACCAACTTAT TCATGAGTCA 1555

CTACTGGCCC AGCCTGGACT TGAACCTGGA ACACAGGGTA GAGCTCAGGC TCTTCAGTAT 1615

CCATCAGAAG ATTTAGGTGT GTGCAGACAT GACCACACTC CCCCTAGCAC TCCATAGCC 1674

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 399 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

Met Ala Ala Arg Pro Gly Leu Leu Trp Leu Leu Gly Leu Ala Leu Cys
1 5 10 15

Val Leu Gly Gly Gly His Leu Ser His Pro Pro His Val Phe Pro Gln
20 25 30

Arg Arg Leu Gly Val Arg Glu Pro Arg Asp Met Gln Arg Glu Ile Arg
35 40 45

Glu Val Leu Gly Leu Ala Gly Arg Pro Arg Ser Arg Ala Pro Val Gly
50 55 60

Ala Ala Gln Gln Pro Ala Ser Ala Pro Leu Phe Met Leu Asp Leu Tyr
65 70 75 80

Arg Ala Met Thr Asp Asp Ser Gly Gly Gly Thr Pro Gln Pro His Leu
85 90 95

Asp Arg Ala Asp Leu Ile Met Ser Phe Val Asn Ile Val Glu Arg Asp
100 105 110

Arg Thr Leu Gly Tyr Gln Glu Pro His Trp Lys Glu Phe His Phe Asp
115 120 125

Leu Thr Gln Ile Pro Ala Gly Glu Ala Val Thr Ala Ala Glu Phe Arg
130 135 140

Ile Tyr Lys Glu Pro Ser Thr His Pro Leu Asn Thr Thr Leu His Ile
145 150 155 160

Ser Met Phe Glu Val Val Gln Glu His Ser Asn Arg Glu Ser Asp Leu
165 170 175

Phe Phe Leu Asp Leu Gln Thr Leu Arg Ser Gly Asp Glu Gly Trp Leu
180 185 190

Val Leu Asp Ile Thr Ala Ala Ser Asp Arg Trp Leu Leu Asn His His
195 200 205

Lys Asp Leu Gly Leu Arg Leu Tyr Val Glu Thr Glu Asp Gly His Ser
210 215 220

Ile Asp Pro Gly Leu Ala Gly Leu Leu Gly Arg Gln Ala Pro Arg Ser
225 230 235 240

Arg Gln Pro Phe Met Val Gly Phe Phe Arg Ala Asn Gln Ser Pro Val
245 250 255

Arg Ala Pro Arg Thr Ala Arg Pro Leu Lys Lys Lys Gln Leu Asn Gln
260 265 270

Ile Asn Gln Leu Pro His Ser Asn Lys His Leu Gly Ile Leu Asp Asp
275 280 285

Gly His Gly Ser His Gly Arg Glu Val Cys Arg Arg His Glu Leu Tyr
290 295 300

Val Ser Phe Arg Asp Leu Gly Trp Leu Asp Ser Val Ile Ala Pro Gln
305 310 315 320

Gly Tyr Ser Ala Tyr Tyr Cys Ala Gly Glu Cys Ile Tyr Pro Leu Asn
325 330 335

Ser Cys Met Asn Ser Thr Asn His Ala Thr Met Gln Ala Leu Val His
340 345 350

Leu Met Lys Pro Asp Ile Ile Pro Lys Val Cys Cys Val Pro Thr Glu
355 360 365

Leu Ser Ala Ile Ser Leu Leu Tyr Tyr Asp Arg Asn Asn Asn Val Ile
370 375 380

Leu Arg Arg Glu Arg Asn Met Val Val Gln Ala Cys Gly Cys His
385 390 395

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1822 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:
(A) ORGANISM: HOMO SAPIENS
(F) TISSUE TYPE: HIPPOCAMPUS

(ix) FEATURE:
(A) NAME/KEY: CDS
(B) LOCATION: 49..1341
(C) IDENTIFICATION METHOD: experimental
(D) OTHER INFORMATION: /function= "OSTEOGENIC PROTEIN"
/product= "hOP1-PP"
/note= "hOP1 cDNA"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

```
GGTGCGGGCC CGGAGCCCGG AGCCCGGGTA GCGCGTAGAG CCGGCGCG ATG CAC GTG      57
                                                      Met His Val
                                                        1

CGC TCA CTG CGA GCT GCG GCG CCG CAC AGC TTC GTG GCG CTC TGG GCA      105
Arg Ser Leu Arg Ala Ala Ala Pro His Ser Phe Val Ala Leu Trp Ala
  5                  10                  15

CCC CTG TTC CTG CTG CGC TCC GCC CTG GCC GAC TTC AGC CTG GAC AAC      153
Pro Leu Phe Leu Leu Arg Ser Ala Leu Ala Asp Phe Ser Leu Asp Asn
 20                  25                  30                  35

GAG GTG CAC TCG AGC TTC ATC CAC CGG CGC CTC CGC AGC CAG GAG CGG      201
Glu Val His Ser Ser Phe Ile His Arg Arg Leu Arg Ser Gln Glu Arg
                 40                  45                  50

CGG GAG ATG CAG CGC GAG ATC CTC TCC ATT TTG GGC TTG CCC CAC CGC      249
Arg Glu Met Gln Arg Glu Ile Leu Ser Ile Leu Gly Leu Pro His Arg
            55                  60                  65
```

```
CCG CGC CCG CAC CTC CAG GGC AAG CAC AAC TCG GCA CCC ATG TTC ATG      297
Pro Arg Pro His Leu Gln Gly Lys His Asn Ser Ala Pro Met Phe Met
        70                  75                  80

CTG GAC CTG TAC AAC GCC ATG GCG GTG GAG GAG GGC GGC GGG CCC GGC      345
Leu Asp Leu Tyr Asn Ala Met Ala Val Glu Glu Gly Gly Gly Pro Gly
    85                  90                  95

GGC CAG GGC TTC TCC TAC CCC TAC AAG GCC GTC TTC AGT ACC CAG GGC      393
Gly Gln Gly Phe Ser Tyr Pro Tyr Lys Ala Val Phe Ser Thr Gln Gly
100                 105                 110                 115

CCC CCT CTG GCC AGC CTG CAA GAT AGC CAT TTC CTC ACC GAC GCC GAC      441
Pro Pro Leu Ala Ser Leu Gln Asp Ser His Phe Leu Thr Asp Ala Asp
                120                 125                 130

ATG GTC ATG AGC TTC GTC AAC CTC GTG GAA CAT GAC AAG GAA TTC TTC      489
Met Val Met Ser Phe Val Asn Leu Val Glu His Asp Lys Glu Phe Phe
            135                 140                 145

CAC CCA CGC TAC CAC CAT CGA GAG TTC CGG TTT GAT CTT TCC AAG ATC      537
His Pro Arg Tyr His His Arg Glu Phe Arg Phe Asp Leu Ser Lys Ile
        150                 155                 160

CCA GAA GGG GAA GCT GTC ACG GCA GCC GAA TTC CGG ATC TAC AAG GAC      585
Pro Glu Gly Glu Ala Val Thr Ala Ala Glu Phe Arg Ile Tyr Lys Asp
    165                 170                 175

TAC ATC CGG GAA CGC TTC GAC AAT GAG ACG TTC CGG ATC AGC GTT TAT      633
Tyr Ile Arg Glu Arg Phe Asp Asn Glu Thr Phe Arg Ile Ser Val Tyr
180                 185                 190                 195

CAG GTG CTC CAG GAG CAC TTG GGC AGG GAA TCG GAT CTC TTC CTG CTC      681
Gln Val Leu Gln Glu His Leu Gly Arg Glu Ser Asp Leu Phe Leu Leu
                200                 205                 210

GAC AGC CGT ACC CTC TGG GCC TCG GAG GAG GGC TGG CTG GTG TTT GAC      729
Asp Ser Arg Thr Leu Trp Ala Ser Glu Glu Gly Trp Leu Val Phe Asp
            215                 220                 225

ATC ACA GCC ACC AGC AAC CAC TGG GTG GTC AAT CCG CGG CAC AAC CTG      777
Ile Thr Ala Thr Ser Asn His Trp Val Val Asn Pro Arg His Asn Leu
        230                 235                 240

GGC CTG CAG CTC TCG GTG GAG ACG CTG GAT GGG CAG AGC ATC AAC CCC      825
Gly Leu Gln Leu Ser Val Glu Thr Leu Asp Gly Gln Ser Ile Asn Pro
    245                 250                 255
```

```
AAG TTG GCG GGC CTG ATT GGG CGG CAC GGG CCC CAG AAC AAG CAG CCC      873
Lys Leu Ala Gly Leu Ile Gly Arg His Gly Pro Gln Asn Lys Gln Pro
260                 265                 270                 275

TTC ATG GTG GCT TTC TTC AAG GCC ACG GAG GTC CAC TTC CGC AGC ATC      921
Phe Met Val Ala Phe Phe Lys Ala Thr Glu Val His Phe Arg Ser Ile
                280                 285                 290

CGG TCC ACG GGG AGC AAA CAG CGC AGC CAG AAC CGC TCC AAG ACG CCC      969
Arg Ser Thr Gly Ser Lys Gln Arg Ser Gln Asn Arg Ser Lys Thr Pro
            295                 300                 305

AAG AAC CAG GAA GCC CTG CGG ATG GCC AAC GTG GCA GAG AAC AGC AGC     1017
Lys Asn Gln Glu Ala Leu Arg Met Ala Asn Val Ala Glu Asn Ser Ser
        310                 315                 320

AGC GAC CAG AGG CAG GCC TGT AAG AAG CAC GAG CTG TAT GTC AGC TTC     1065
Ser Asp Gln Arg Gln Ala Cys Lys Lys His Glu Leu Tyr Val Ser Phe
    325                 330                 335

CGA GAC CTG GGC TGG CAG GAC TGG ATC ATC GCG CCT GAA GGC TAC GCC     1113
Arg Asp Leu Gly Trp Gln Asp Trp Ile Ile Ala Pro Glu Gly Tyr Ala
340                 345                 350                 355

GCC TAC TAC TGT GAG GGG GAG TGT GCC TTC CCT CTG AAC TCC TAC ATG     1161
Ala Tyr Tyr Cys Glu Gly Glu Cys Ala Phe Pro Leu Asn Ser Tyr Met
                360                 365                 370

AAC GCC ACC AAC CAC GCC ATC GTG CAG ACG CTG GTC CAC TTC ATC AAC     1209
Asn Ala Thr Asn His Ala Ile Val Gln Thr Leu Val His Phe Ile Asn
            375                 380                 385

CCG GAA ACG GTG CCC AAG CCC TGC TGT GCG CCC ACG CAG CTC AAT GCC     1257
Pro Glu Thr Val Pro Lys Pro Cys Cys Ala Pro Thr Gln Leu Asn Ala
        390                 395                 400

ATC TCC GTC CTC TAC TTC GAT GAC AGC TCC AAC GTC ATC CTG AAG AAA     1305
Ile Ser Val Leu Tyr Phe Asp Asp Ser Ser Asn Val Ile Leu Lys Lys
    405                 410                 415

TAC AGA AAC ATG GTG GTC CGG GCC TGT GGC TGC CAC TAGCTCCTCC          1351
Tyr Arg Asn Met Val Val Arg Ala Cys Gly Cys His
420                 425                 430

GAGAATTCAG ACCCTTTGGG GCCAAGTTTT TCTGGATCCT CCATTGCTCG CCTTGGCCAG   1411

GAACCAGCAG ACCAACTGCC TTTTGTGAGA CCTTCCCCTC CCTATCCCCA ACTTTAAAGG   1471
```

TGTGAGAGTA TTAGGAAACA TGAGCAGCAT ATGGCTTTTG ATCAGTTTTT CAGTGGCAGC 1531

ATCCAATGAA CAAGATCCTA CAAGCTGTGC AGGCAAAACC TAGCAGGAAA AAAAAACAAC 1591

GCATAAAGAA AAATGGCCGG GCCAGGTCAT TGGCTGGGAA GTCTCAGCCA TGCACGGACT 1651

CGTTTCCAGA GGTAATTATG AGCGCCTACC AGCCAGGCCA CCCAGCCGTG GGAGGAAGGG 1711

GGCGTGGCAA GGGGTGGGCA CATTGGTGTC TGTGCGAAAG GAAAATTGAC CCGGAAGTTC 1771

CTGTAATAAA TGTCACAATA AAACGAATGA ATGAAAAAAA AAAAAAAAAA A 1822

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 431 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

Met His Val Arg Ser Leu Arg Ala Ala Ala Pro His Ser Phe Val Ala
1 5 10 15

Leu Trp Ala Pro Leu Phe Leu Leu Arg Ser Ala Leu Ala Asp Phe Ser
20 25 30

Leu Asp Asn Glu Val His Ser Ser Phe Ile His Arg Arg Leu Arg Ser
35 40 45

Gln Glu Arg Arg Glu Met Gln Arg Glu Ile Leu Ser Ile Leu Gly Leu
50 55 60

Pro His Arg Pro Arg Pro His Leu Gln Gly Lys His Asn Ser Ala Pro
65 70 75 80

Met Phe Met Leu Asp Leu Tyr Asn Ala Met Ala Val Glu Glu Gly Gly
85 90 95

Gly Pro Gly Gly Gln Gly Phe Ser Tyr Pro Tyr Lys Ala Val Phe Ser
100 105 110

Thr Gln Gly Pro Pro Leu Ala Ser Leu Gln Asp Ser His Phe Leu Thr
115 120 125

Asp Ala Asp Met Val Met Ser Phe Val Asn Leu Val Glu His Asp Lys
130 135 140

Glu Phe Phe His Pro Arg Tyr His His Arg Glu Phe Arg Phe Asp Leu
145 150 155 160

Ser Lys Ile Pro Glu Gly Glu Ala Val Thr Ala Ala Glu Phe Arg Ile
165 170 175

Tyr Lys Asp Tyr Ile Arg Glu Arg Phe Asp Asn Glu Thr Phe Arg Ile
180 185 190

Ser Val Tyr Gln Val Leu Gln Glu His Leu Gly Arg Glu Ser Asp Leu
195 200 205

Phe Leu Leu Asp Ser Arg Thr Leu Trp Ala Ser Glu Glu Gly Trp Leu
210 215 220

Val Phe Asp Ile Thr Ala Thr Ser Asn His Trp Val Val Asn Pro Arg
225 230 235 240

His Asn Leu Gly Leu Gln Leu Ser Val Glu Thr Leu Asp Gly Gln Ser
245 250 255

Ile Asn Pro Lys Leu Ala Gly Leu Ile Gly Arg His Gly Pro Gln Asn
260 265 270

Lys Gln Pro Phe Met Val Ala Phe Phe Lys Ala Thr Glu Val His Phe
275 280 285

Arg Ser Ile Arg Ser Thr Gly Ser Lys Gln Arg Ser Gln Asn Arg Ser
290 295 300

Lys Thr Pro Lys Asn Gln Glu Ala Leu Arg Met Ala Asn Val Ala Glu
305 310 315 320

Asn Ser Ser Ser Asp Gln Arg Gln Ala Cys Lys Lys His Glu Leu Tyr
325 330 335

Val Ser Phe Arg Asp Leu Gly Trp Gln Asp Trp Ile Ile Ala Pro Glu
340 345 350

Gly Tyr Ala Ala Tyr Tyr Cys Glu Gly Glu Cys Ala Phe Pro Leu Asn
355 360 365

Ser Tyr Met Asn Ala Thr Asn His Ala Ile Val Gln Thr Leu Val His
370 375 380

Phe Ile Asn Pro Glu Thr Val Pro Lys Pro Cys Cys Ala Pro Thr Gln
385 390 395 400

Leu Asn Ala Ile Ser Val Leu Tyr Phe Asp Asp Ser Ser Asn Val Ile
405 410 415

Leu Lys Lys Tyr Arg Asn Met Val Val Arg Ala Cys Gly Cys His
420 425 430

(2) INFORMATION FOR SEQ ID NO:5:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1873 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:
(A) ORGANISM: MURIDAE
(F) TISSUE TYPE: EMBRYO

(ix) FEATURE:
(A) NAME/KEY: CDS
(B) LOCATION: 104..1393
(D) OTHER INFORMATION: /function= "OSTEOGENIC PROTEIN"
/product= "MOP1-PP"
/note= "MOP1 (cDNA)"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

CTGCAGCAAG TGACCTCGGG TCGTGGACCG CTGCCCTGCC CCCTCCGCTG CCACCTGGGG 60

CGGCGCGGGC CCGGTGCCCC GGATCGCGCG TAGAGCCGGC GCG ATG CAC GTG CGC 115
Met His Val Arg
1

TCG CTG CGC GCT GCG GCG CCA CAC AGC TTC GTG GCG CTC TGG GCG CCT 163
Ser Leu Arg Ala Ala Ala Pro His Ser Phe Val Ala Leu Trp Ala Pro
5 10 15 20

CTG TTC TTG CTG CGC TCC GCC CTG GCC GAT TTC AGC CTG GAC AAC GAG 211
Leu Phe Leu Leu Arg Ser Ala Leu Ala Asp Phe Ser Leu Asp Asn Glu
25 30 35

```
GTG CAC TCC AGC TTC ATC CAC CGG CGC CTC CGC AGC CAG GAG CGG CGG      259
Val His Ser Ser Phe Ile His Arg Arg Leu Arg Ser Gln Glu Arg Arg
          40                  45                  50

GAG ATG CAG CGG GAG ATC CTG TCC ATC TTA GGG TTG CCC CAT CGC CCG      307
Glu Met Gln Arg Glu Ile Leu Ser Ile Leu Gly Leu Pro His Arg Pro
     55                  60                  65

CGC CCG CAC CTC CAG GGA AAG CAT AAT TCG GCG CCC ATG TTC ATG TTG      355
Arg Pro His Leu Gln Gly Lys His Asn Ser Ala Pro Met Phe Met Leu
70                  75                  80

GAC CTG TAC AAC GCC ATG GCG GTG GAG GAG AGC GGG CCG GAC GGA CAG      403
Asp Leu Tyr Asn Ala Met Ala Val Glu Glu Ser Gly Pro Asp Gly Gln
85                  90                  95                  100

GGC TTC TCC TAC CCC TAC AAG GCC GTC TTC AGT ACC CAG GGC CCC CCT      451
Gly Phe Ser Tyr Pro Tyr Lys Ala Val Phe Ser Thr Gln Gly Pro Pro
               105                 110                 115

TTA GCC AGC CTG CAG GAC AGC CAT TTC CTC ACT GAC GCC GAC ATG GTC      499
Leu Ala Ser Leu Gln Asp Ser His Phe Leu Thr Asp Ala Asp Met Val
          120                 125                 130

ATG AGC TTC GTC AAC CTA GTG GAA CAT GAC AAA GAA TTC TTC CAC CCT      547
Met Ser Phe Val Asn Leu Val Glu His Asp Lys Glu Phe Phe His Pro
     135                 140                 145

CGA TAC CAC CAT CGG GAG TTC CGG TTT GAT CTT TCC AAG ATC CCC GAG      595
Arg Tyr His His Arg Glu Phe Arg Phe Asp Leu Ser Lys Ile Pro Glu
150                 155                 160

GGC GAA CGG GTG ACC GCA GCC GAA TTC AGG ATC TAT AAG GAC TAC ATC      643
Gly Glu Arg Val Thr Ala Ala Glu Phe Arg Ile Tyr Lys Asp Tyr Ile
165                 170                 175                 180

CGG GAG CGA TTT GAC AAC GAG ACC TTC CAG ATC ACA GTC TAT CAG GTG      691
Arg Glu Arg Phe Asp Asn Glu Thr Phe Gln Ile Thr Val Tyr Gln Val
               185                 190                 195

CTC CAG GAG CAC TCA GGC AGG GAG TCG GAC CTC TTC TTG CTG GAC AGC      739
Leu Gln Glu His Ser Gly Arg Glu Ser Asp Leu Phe Leu Leu Asp Ser
          200                 205                 210

CGC ACC ATC TGG GCT TCT GAG GAG GGC TGG TTG GTG TTT GAT ATC ACA      787
Arg Thr Ile Trp Ala Ser Glu Glu Gly Trp Leu Val Phe Asp Ile Thr
     215                 220                 225
```

```
GCC ACC AGC AAC CAC TGG GTG GTC AAC CCT CGG CAC AAC CTG GGC TTA      835
Ala Thr Ser Asn His Trp Val Val Asn Pro Arg His Asn Leu Gly Leu
    230                 235                 240

CAG CTC TCT GTG GAG ACC CTG GAT GGG CAG AGC ATC AAC CCC AAG TTG      883
Gln Leu Ser Val Glu Thr Leu Asp Gly Gln Ser Ile Asn Pro Lys Leu
245                 250                 255                 260

GCA GGC CTG ATT GGA CGG CAT GGA CCC CAG AAC AAG CAA CCC TTC ATG      931
Ala Gly Leu Ile Gly Arg His Gly Pro Gln Asn Lys Gln Pro Phe Met
                265                 270                 275

GTG GCC TTC TTC AAG GCC ACG GAA GTC CAT CTC CGT AGT ATC CGG TCC      979
Val Ala Phe Phe Lys Ala Thr Glu Val His Leu Arg Ser Ile Arg Ser
            280                 285                 290

ACG GGG GGC AAG CAG CGC AGC CAG AAT CGC TCC AAG ACG CCA AAG AAC      1027
Thr Gly Gly Lys Gln Arg Ser Gln Asn Arg Ser Lys Thr Pro Lys Asn
        295                 300                 305

CAA GAG GCC CTG AGG ATG GCC AGT GTG GCA GAA AAC AGC AGC AGT GAC      1075
Gln Glu Ala Leu Arg Met Ala Ser Val Ala Glu Asn Ser Ser Ser Asp
    310                 315                 320

CAG AGG CAG GCC TGC AAG AAA CAT GAG CTG TAC GTC AGC TTC CGA GAC      1123
Gln Arg Gln Ala Cys Lys Lys His Glu Leu Tyr Val Ser Phe Arg Asp
325                 330                 335                 340

CTT GGC TGG CAG GAC TGG ATC ATT GCA CCT GAA GGC TAT GCT GCC TAC      1171
Leu Gly Trp Gln Asp Trp Ile Ile Ala Pro Glu Gly Tyr Ala Ala Tyr
                345                 350                 355

TAC TGT GAG GGA GAG TGC GCC TTC CCT CTG AAC TCC TAC ATG AAC GCC      1219
Tyr Cys Glu Gly Glu Cys Ala Phe Pro Leu Asn Ser Tyr Met Asn Ala
            360                 365                 370

ACC AAC CAC GCC ATC GTC CAG ACA CTG GTT CAC TTC ATC AAC CCA GAC      1267
Thr Asn His Ala Ile Val Gln Thr Leu Val His Phe Ile Asn Pro Asp
        375                 380                 385

ACA GTA CCC AAG CCC TGC TGT GCG CCC ACC CAG CTC AAC GCC ATC TCT      1315
Thr Val Pro Lys Pro Cys Cys Ala Pro Thr Gln Leu Asn Ala Ile Ser
    390                 395                 400

GTC CTC TAC TTC GAC GAC AGC TCT AAT GTC GAC CTG AAG AAG TAC AGA      1363
Val Leu Tyr Phe Asp Asp Ser Ser Asn Val Asp Leu Lys Lys Tyr Arg
405                 410                 415                 420
```

AAC ATG GTG GTC CGG GCC TGT GGC TGC CAC TAGCTCTTCC TGAGACCCTG 1413
Asn Met Val Val Arg Ala Cys Gly Cys His
425 430

ACCTTTGCGG GGCCACACCT TTCCAAATCT TCGATGTCTC ACCATCTAAG TCTCTCACTG 1473

CCCACCTTGG CGAGGAGAAC AGACCAACCT CTCCTGAGCC TTCCCTCACC TCCCAACCGG 1533

AAGCATGTAA GGGTTCCAGA AACCTGAGCG TGCAGCAGCT GATGAGCGCC CTTTCCTTCT 1593

GGCACGTGAC GGACAAGATC CTACCAGCTA CCACAGCAAA CGCCTAAGAG CAGGAAAAAT 1653

GTCTGCCAGG AAAGTGTCCA GTGTCCACAT GGCCCCTGGC GCTCTGAGTC TTTGAGGAGT 1713

AATCGCAAGC CTCGTTCAGC TGCAGCAGAA GGAAGGGCTT AGCCAGGGTG GGCGCTGGCG 1773

TCTGTGTTGA AGGGAAACCA AGCAGAAGCC ACTGTAATGA TATGTCACAA TAAAACCCAT 1833

GAATGAAAAA AAAAAAAAAA AAAAAAAAAA AAAAGAATTC 1873

(2) INFORMATION FOR SEQ ID NO:6:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 430 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

Met His Val Arg Ser Leu Arg Ala Ala Ala Pro His Ser Phe Val Ala
1 5 10 15

Leu Trp Ala Pro Leu Phe Leu Leu Arg Ser Ala Leu Ala Asp Phe Ser
20 25 30

Leu Asp Asn Glu Val His Ser Ser Phe Ile His Arg Arg Leu Arg Ser
35 40 45

Gln Glu Arg Arg Glu Met Gln Arg Glu Ile Leu Ser Ile Leu Gly Leu
50 55 60

Pro His Arg Pro Arg Pro His Leu Gln Gly Lys His Asn Ser Ala Pro
65 70 75 80

Met Phe Met Leu Asp Leu Tyr Asn Ala Met Ala Val Glu Glu Ser Gly
85 90 95

Pro Asp Gly Gln Gly Phe Ser Tyr Pro Tyr Lys Ala Val Phe Ser Thr
100 105 110

Gln Gly Pro Pro Leu Ala Ser Leu Gln Asp Ser His Phe Leu Thr Asp
115 120 125

Ala Asp Met Val Met Ser Phe Val Asn Leu Val Glu His Asp Lys Glu
130 135 140

Phe Phe His Pro Arg Tyr His His Arg Glu Phe Arg Phe Asp Leu Ser
145 150 155 160

Lys Ile Pro Glu Gly Glu Arg Val Thr Ala Ala Glu Phe Arg Ile Tyr
165 170 175

Lys Asp Tyr Ile Arg Glu Arg Phe Asp Asn Glu Thr Phe Gln Ile Thr
180 185 190

Val Tyr Gln Val Leu Gln Glu His Ser Gly Arg Glu Ser Asp Leu Phe
195 200 205

Leu Leu Asp Ser Arg Thr Ile Trp Ala Ser Glu Glu Gly Trp Leu Val
210 215 220

Phe Asp Ile Thr Ala Thr Ser Asn His Trp Val Val Asn Pro Arg His
225 230 235 240

Asn Leu Gly Leu Gln Leu Ser Val Glu Thr Leu Asp Gly Gln Ser Ile
245 250 255

Asn Pro Lys Leu Ala Gly Leu Ile Gly Arg His Gly Pro Gln Asn Lys
260 265 270

Gln Pro Phe Met Val Ala Phe Phe Lys Ala Thr Glu Val His Leu Arg
275 280 285

Ser Ile Arg Ser Thr Gly Gly Lys Gln Arg Ser Gln Asn Arg Ser Lys
290 295 300

Thr Pro Lys Asn Gln Glu Ala Leu Arg Met Ala Ser Val Ala Glu Asn
305 310 315 320

Ser Ser Ser Asp Gln Arg Gln Ala Cys Lys Lys His Glu Leu Tyr Val
325 330 335

Ser Phe Arg Asp Leu Gly Trp Gln Asp Trp Ile Ile Ala Pro Glu Gly
340 345 350

Tyr Ala Ala Tyr Tyr Cys Glu Gly Glu Cys Ala Phe Pro Leu Asn Ser
355 360 365

Tyr Met Asn Ala Thr Asn His Ala Ile Val Gln Thr Leu Val His Phe
370 375 380

Ile Asn Pro Asp Thr Val Pro Lys Pro Cys Cys Ala Pro Thr Gln Leu
385 390 395 400

Asn Ala Ile Ser Val Leu Tyr Phe Asp Asp Ser Ser Asn Val Asp Leu
405 410 415

Lys Lys Tyr Arg Asn Met Val Val Arg Ala Cys Gly Cys His
420 425 430

(2) INFORMATION FOR SEQ ID NO:7:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1723 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo sapiens
(F) TISSUE TYPE: HIPPOCAMPUS

(ix) FEATURE:
(A) NAME/KEY: CDS
(B) LOCATION: 490..1696
(D) OTHER INFORMATION: /function= "OSTEOGENIC PROTEIN"
/product= "hOP2-PP"
/note= "hOP2 (cDNA)"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

GGCGCCGGCA GAGCAGGAGT GGCTGGAGGA GCTGTGGTTG GAGCAGGAGG TGGCACGGCA 60

GGGCTGGAGG GCTCCCTATG AGTGGCGGAG ACGGCCCAGG AGGCGCTGGA GCAACAGCTC 120

CCACACCGCA CCAAGCGGTG GCTGCAGGAG CTCGCCCATC GCCCCTGCGC TGCTCGGACC 180

```
GCGGCCACAG CCGGACTGGC GGGTACGGCG GCGACAGAGG CATTGGCCGA GAGTCCCAGT      240

CCGCAGAGTA GCCCCGGCCT CGAGGCGGTG GCGTCCCGGT CCTCTCCGTC CAGGAGCCAG      300

GACAGGTGTC GCGCGGCGGG GCTCCAGGGA CCGCGCCTGA GGCCGGCTGC CCGCCCGTCC      360

CGCCCCGCCC CGCCGCCCGC CGCCCGCCGA GCCCAGCCTC CTTGCCGTCG GGGCGTCCCC      420

AGGCCCTGGG TCGGCCGCGG AGCCGATGCG CGCCCGCTGA GCGCCCCAGC TGAGCGCCCC      480

CGGCCTGCC ATG ACC GCG CTC CCC GGC CCG CTC TGG CTC CTG GGC CTG           528
          Met Thr Ala Leu Pro Gly Pro Leu Trp Leu Leu Gly Leu
            1               5                  10

GCG CTA TGC GCG CTG GGC GGG GGC GGC CCC GGC CTG CGA CCC CCG CCC        576
Ala Leu Cys Ala Leu Gly Gly Gly Gly Pro Gly Leu Arg Pro Pro Pro
     15                  20                  25

GGC TGT CCC CAG CGA CGT CTG GGC GCG CGC GAG CGC CGG GAC GTG CAG        624
Gly Cys Pro Gln Arg Arg Leu Gly Ala Arg Glu Arg Arg Asp Val Gln
 30                  35                  40                  45

CGC GAG ATC CTG GCG GTG CTC GGG CTG CCT GGG CGG CCC CGG CCC CGC        672
Arg Glu Ile Leu Ala Val Leu Gly Leu Pro Gly Arg Pro Arg Pro Arg
                 50                  55                  60

GCG CCA CCC GCC GCC TCC CGG CTG CCC GCG TCC GCG CCG CTC TTC ATG        720
Ala Pro Pro Ala Ala Ser Arg Leu Pro Ala Ser Ala Pro Leu Phe Met
            65                  70                  75

CTG GAC CTG TAC CAC GCC ATG GCC GGC GAC GAC GAC GAG GAC GGC GCG        768
Leu Asp Leu Tyr His Ala Met Ala Gly Asp Asp Asp Glu Asp Gly Ala
        80                  85                  90

CCC GCG GAG CGG CGC CTG GGC CGC GCC GAC CTG GTC ATG AGC TTC GTT        816
Pro Ala Glu Arg Arg Leu Gly Arg Ala Asp Leu Val Met Ser Phe Val
     95                  100                 105

AAC ATG GTG GAG CGA GAC CGT GCC CTG GGC CAC CAG GAG CCC CAT TGG        864
Asn Met Val Glu Arg Asp Arg Ala Leu Gly His Gln Glu Pro His Trp
110                 115                 120                 125

AAG GAG TTC CGC TTT GAC CTG ACC CAG ATC CCG GCT GGG GAG GCG GTC        912
Lys Glu Phe Arg Phe Asp Leu Thr Gln Ile Pro Ala Gly Glu Ala Val
                130                 135                 140

ACA GCT GCG GAG TTC CGG ATT TAC AAG GTG CCC AGC ATC CAC CTG CTC        960
Thr Ala Ala Glu Phe Arg Ile Tyr Lys Val Pro Ser Ile His Leu Leu
            145                 150                 155
```

```
AAC AGG ACC CTC CAC GTC AGC ATG TTC CAG GTG GTC CAG GAG CAG TCC      1008
Asn Arg Thr Leu His Val Ser Met Phe Gln Val Val Gln Glu Gln Ser
        160                 165                 170

AAC AGG GAG TCT GAC TTG TTC TTT TTG GAT CTT CAG ACG CTC CGA GCT      1056
Asn Arg Glu Ser Asp Leu Phe Phe Leu Asp Leu Gln Thr Leu Arg Ala
    175                 180                 185

GGA GAC GAG GGC TGG CTG GTG CTG GAT GTC ACA GCA GCC AGT GAC TGC      1104
Gly Asp Glu Gly Trp Leu Val Leu Asp Val Thr Ala Ala Ser Asp Cys
190                 195                 200                 205

TGG TTG CTG AAG CGT CAC AAG GAC CTG GGA CTC CGC CTC TAT GTG GAG      1152
Trp Leu Leu Lys Arg His Lys Asp Leu Gly Leu Arg Leu Tyr Val Glu
                210                 215                 220

ACT GAG GAC GGG CAC AGC GTG GAT CCT GGC CTG GCC GGC CTG CTG GGT      1200
Thr Glu Asp Gly His Ser Val Asp Pro Gly Leu Ala Gly Leu Leu Gly
            225                 230                 235

CAA CGG GCC CCA CGC TCC CAA CAG CCT TTC GTG GTC ACT TTC TTC AGG      1248
Gln Arg Ala Pro Arg Ser Gln Gln Pro Phe Val Val Thr Phe Phe Arg
        240                 245                 250

GCC AGT CCG AGT CCC ATC CGC ACC CCT CGG GCA GTG AGG CCA CTG AGG      1296
Ala Ser Pro Ser Pro Ile Arg Thr Pro Arg Ala Val Arg Pro Leu Arg
    255                 260                 265

AGG AGG CAG CCG AAG AAA AGC AAC GAG CTG CCG CAG GCC AAC CGA CTC      1344
Arg Arg Gln Pro Lys Lys Ser Asn Glu Leu Pro Gln Ala Asn Arg Leu
270                 275                 280                 285

CCA GGG ATC TTT GAT GAC GTC CAC GGC TCC CAC GGC CGG CAG GTC TGC      1392
Pro Gly Ile Phe Asp Asp Val His Gly Ser His Gly Arg Gln Val Cys
                290                 295                 300

CGT CGG CAC GAG CTC TAC GTC AGC TTC CAG GAC CTC GGC TGG CTG GAC      1440
Arg Arg His Glu Leu Tyr Val Ser Phe Gln Asp Leu Gly Trp Leu Asp
            305                 310                 315

TGG GTC ATC GCT CCC CAA GGC TAC TCG GCC TAT TAC TGT GAG GGG GAG      1488
Trp Val Ile Ala Pro Gln Gly Tyr Ser Ala Tyr Tyr Cys Glu Gly Glu
        320                 325                 330

TGC TCC TTC CCA CTG GAC TCC TGC ATG AAT GCC ACC AAC CAC GCC ATC      1536
Cys Ser Phe Pro Leu Asp Ser Cys Met Asn Ala Thr Asn His Ala Ile
    335                 340                 345
```

```
CTG CAG TCC CTG GTG CAC CTG ATG AAG CCA AAC GCA GTC CCC AAG GCG      1584
Leu Gln Ser Leu Val His Leu Met Lys Pro Asn Ala Val Pro Lys Ala
350                 355                 360                 365

TGC TGT GCA CCC ACC AAG CTG AGC GCC ACC TCT GTG CTC TAC TAT GAC      1632
Cys Cys Ala Pro Thr Lys Leu Ser Ala Thr Ser Val Leu Tyr Tyr Asp
                370                 375                 380

AGC AGC AAC AAC GTC ATC CTG CGC AAA GCC CGC AAC ATG GTG GTC AAG      1680
Ser Ser Asn Asn Val Ile Leu Arg Lys Ala Arg Asn Met Val Val Lys
            385                 390                 395

GCC TGC GGC TGC CAC T GAGTCAGCCC GCCCAGCCCT ACTGCAG                  1723
Ala Cys Gly Cys His
        400
```

(2) INFORMATION FOR SEQ ID NO:8:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 402 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

```
Met Thr Ala Leu Pro Gly Pro Leu Trp Leu Leu Gly Leu Ala Leu Cys
 1               5                  10                  15

Ala Leu Gly Gly Gly Gly Pro Gly Leu Arg Pro Pro Pro Gly Cys Pro
            20                  25                  30

Gln Arg Arg Leu Gly Ala Arg Glu Arg Arg Asp Val Gln Arg Glu Ile
        35                  40                  45

Leu Ala Val Leu Gly Leu Pro Gly Arg Pro Arg Pro Arg Ala Pro Pro
    50                  55                  60

Ala Ala Ser Arg Leu Pro Ala Ser Ala Pro Leu Phe Met Leu Asp Leu
65                  70                  75                  80

Tyr His Ala Met Ala Gly Asp Asp Asp Glu Asp Gly Ala Pro Ala Glu
                85                  90                  95

Arg Arg Leu Gly Arg Ala Asp Leu Val Met Ser Phe Val Asn Met Val
            100                 105                 110
```

Glu Arg Asp Arg Ala Leu Gly His Gln Glu Pro His Trp Lys Glu Phe
115 120 125

Arg Phe Asp Leu Thr Gln Ile Pro Ala Gly Glu Ala Val Thr Ala Ala
130 135 140

Glu Phe Arg Ile Tyr Lys Val Pro Ser Ile His Leu Leu Asn Arg Thr
145 150 155 160

Leu His Val Ser Met Phe Gln Val Val Gln Glu Gln Ser Asn Arg Glu
165 170 175

Ser Asp Leu Phe Phe Leu Asp Leu Gln Thr Leu Arg Ala Gly Asp Glu
180 185 190

Gly Trp Leu Val Leu Asp Val Thr Ala Ala Ser Asp Cys Trp Leu Leu
195 200 205

Lys Arg His Lys Asp Leu Gly Leu Arg Leu Tyr Val Glu Thr Glu Asp
210 215 220

Gly His Ser Val Asp Pro Gly Leu Ala Gly Leu Leu Gly Gln Arg Ala
225 230 235 240

Pro Arg Ser Gln Gln Pro Phe Val Val Thr Phe Phe Arg Ala Ser Pro
245 250 255

Ser Pro Ile Arg Thr Pro Arg Ala Val Arg Pro Leu Arg Arg Arg Gln
260 265 270

Pro Lys Lys Ser Asn Glu Leu Pro Gln Ala Asn Arg Leu Pro Gly Ile
275 280 285

Phe Asp Asp Val His Gly Ser His Gly Arg Gln Val Cys Arg Arg His
290 295 300

Glu Leu Tyr Val Ser Phe Gln Asp Leu Gly Trp Leu Asp Trp Val Ile
305 310 315 320

Ala Pro Gln Gly Tyr Ser Ala Tyr Tyr Cys Glu Gly Glu Cys Ser Phe
325 330 335

Pro Leu Asp Ser Cys Met Asn Ala Thr Asn His Ala Ile Leu Gln Ser
340 345 350

Leu Val His Leu Met Lys Pro Asn Ala Val Pro Lys Ala Cys Cys Ala
355 360 365

```
Pro Thr Lys Leu Ser Ala Thr Ser Val Leu Tyr Tyr Asp Ser Ser Asn
    370                 375                 380

Asn Val Ile Leu Arg Lys Ala Arg Asn Met Val Val Lys Ala Cys Gly
385                 390                 395                 400

Cys His
```

(2) INFORMATION FOR SEQ ID NO:9:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1926 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(vi) ORIGINAL SOURCE:
(A) ORGANISM: MURIDAE
(F) TISSUE TYPE: EMBRYO

(ix) FEATURE:
(A) NAME/KEY: CDS
(B) LOCATION: 93..1289
(D) OTHER INFORMATION: /function= "OSTEOGENIC PROTEIN"
/product= "mOP2-PP"
/note= "mOP2 cDNA"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

```
GCCAGGCACA GGTGCGCCGT CTGGTCCTCC CCGTCTGGCG TCAGCCGAGC CCGACCAGCT      60

ACCAGTGGAT GCGCGCCGGC TGAAAGTCCG AG ATG GCT ATG CGT CCC GGG CCA      113
                                    Met Ala Met Arg Pro Gly Pro
                                      1               5

CTC TGG CTA TTG GGC CTT GCT CTG TGC GCG CTG GGA GGC GGC CAC GGT      161
Leu Trp Leu Leu Gly Leu Ala Leu Cys Ala Leu Gly Gly Gly His Gly
         10                  15                  20

CCG CGT CCC CCG CAC ACC TGT CCC CAG CGT CGC CTG GGA GCG CGC GAG      209
Pro Arg Pro Pro His Thr Cys Pro Gln Arg Arg Leu Gly Ala Arg Glu
     25                  30                  35

CGC CGC GAC ATG CAG CGT GAA ATC CTG GCG GTG CTC GGG CTA CCG GGA      257
Arg Arg Asp Met Gln Arg Glu Ile Leu Ala Val Leu Gly Leu Pro Gly
 40                  45                  50                  55
```

```
CGG CCC CGA CCC CGT GCA CAA CCC GCC GCT GCC CGG CAG CCA GCG TCC      305
Arg Pro Arg Pro Arg Ala Gln Pro Ala Ala Ala Arg Gln Pro Ala Ser
                60                  65                  70

GCG CCC CTC TTC ATG TTG GAC CTA TAC CAC GCC ATG ACC GAT GAC GAC      353
Ala Pro Leu Phe Met Leu Asp Leu Tyr His Ala Met Thr Asp Asp Asp
            75                  80                  85

GAC GGC GGG CCA CCA CAG GCT CAC TTA GGC CGT GCC GAC CTG GTC ATG      401
Asp Gly Gly Pro Pro Gln Ala His Leu Gly Arg Ala Asp Leu Val Met
        90                  95                  100

AGC TTC GTC AAC ATG GTG GAA CGC GAC CGT ACC CTG GGC TAC CAG GAG      449
Ser Phe Val Asn Met Val Glu Arg Asp Arg Thr Leu Gly Tyr Gln Glu
    105                 110                 115

CCA CAC TGG AAG GAA TTC CAC TTT GAC CTA ACC CAG ATC CCT GCT GGG      497
Pro His Trp Lys Glu Phe His Phe Asp Leu Thr Gln Ile Pro Ala Gly
120                 125                 130                 135

GAG GCT GTC ACA GCT GCT GAG TTC CGG ATC TAC AAA GAA CCC AGC ACC      545
Glu Ala Val Thr Ala Ala Glu Phe Arg Ile Tyr Lys Glu Pro Ser Thr
                140                 145                 150

CAC CCG CTC AAC ACA ACC CTC CAC ATC AGC ATG TTC GAA GTG GTC CAA      593
His Pro Leu Asn Thr Thr Leu His Ile Ser Met Phe Glu Val Val Gln
            155                 160                 165

GAG CAC TCC AAC AGG GAG TCT GAC TTG TTC TTT TTG GAT CTT CAG ACG      641
Glu His Ser Asn Arg Glu Ser Asp Leu Phe Phe Leu Asp Leu Gln Thr
        170                 175                 180

CTC CGA TCT GGG GAC GAG GGC TGG CTG GTG CTG GAC ATC ACA GCA GCC      689
Leu Arg Ser Gly Asp Glu Gly Trp Leu Val Leu Asp Ile Thr Ala Ala
    185                 190                 195

AGT GAC CGA TGG CTG CTG AAC CAT CAC AAG GAC CTG GGA CTC CGC CTC      737
Ser Asp Arg Trp Leu Leu Asn His His Lys Asp Leu Gly Leu Arg Leu
200                 205                 210                 215

TAT GTG GAA ACC GCG GAT GGG CAC AGC ATG GAT CCT GGC CTG GCT GGT      785
Tyr Val Glu Thr Ala Asp Gly His Ser Met Asp Pro Gly Leu Ala Gly
                220                 225                 230

CTG CTT GGA CGA CAA GCA CCA CGC TCC AGA CAG CCT TTC ATG GTA ACC      833
Leu Leu Gly Arg Gln Ala Pro Arg Ser Arg Gln Pro Phe Met Val Thr
            235                 240                 245
```

```
TTC TTC AGG GCC AGC CAG AGT CCT GTG CGG GCC CCT CGG GCA GCG AGA      881
Phe Phe Arg Ala Ser Gln Ser Pro Val Arg Ala Pro Arg Ala Ala Arg
        250                 255                 260

CCA CTG AAG AGG AGG CAG CCA AAG AAA ACG AAC GAG CTT CCG CAC CCC      929
Pro Leu Lys Arg Arg Gln Pro Lys Lys Thr Asn Glu Leu Pro His Pro
    265                 270                 275

AAC AAA CTC CCA GGG ATC TTT GAT GAT GGC CAC GGT TCC CGC GGC AGA      977
Asn Lys Leu Pro Gly Ile Phe Asp Asp Gly His Gly Ser Arg Gly Arg
280                 285                 290                 295

GAG GTT TGC CGC AGG CAT GAG CTC TAC GTC AGC TTC CGT GAC CTT GGC     1025
Glu Val Cys Arg Arg His Glu Leu Tyr Val Ser Phe Arg Asp Leu Gly
                300                 305                 310

TGG CTG GAC TGG GTC ATC GCC CCC CAG GGC TAC TCT GCC TAT TAC TGT     1073
Trp Leu Asp Trp Val Ile Ala Pro Gln Gly Tyr Ser Ala Tyr Tyr Cys
            315                 320                 325

GAG GGG GAG TGT GCT TTC CCA CTG GAC TCC TGT ATG AAC GCC ACC AAC     1121
Glu Gly Glu Cys Ala Phe Pro Leu Asp Ser Cys Met Asn Ala Thr Asn
        330                 335                 340

CAT GCC ATC TTG CAG TCT CTG GTG CAC CTG ATG AAG CCA GAT GTT GTC     1169
His Ala Ile Leu Gln Ser Leu Val His Leu Met Lys Pro Asp Val Val
    345                 350                 355

CCC AAG GCA TGC TGT GCA CCC ACC AAA CTG AGT GCC ACC TCT GTG CTG     1217
Pro Lys Ala Cys Cys Ala Pro Thr Lys Leu Ser Ala Thr Ser Val Leu
360                 365                 370                 375

TAC TAT GAC AGC AGC AAC AAT GTC ATC CTG CGT AAA CAC CGT AAC ATG     1265
Tyr Tyr Asp Ser Ser Asn Asn Val Ile Leu Arg Lys His Arg Asn Met
                380                 385                 390

GTG GTC AAG GCC TGT GGC TGC CAC TGAGGCCCCG CCCAGCATCC TGCTTCTACT    1319
Val Val Lys Ala Cys Gly Cys His
            395

ACCTTACCAT CTGGCCGGGC CCCTCTCCAG AGGCAGAAAC CCTTCTATGT TATCATAGCT   1379

CAGACAGGGG CAATGGGAGG CCCTTCACTT CCCCTGGCCA CTTCCTGCTA AAATTCTGGT   1439

CTTTCCCAGT TCCTCTGTCC TTCATGGGGT TTCGGGGCTA TCACCCCGCC CTCTCCATCC   1499

TCCTACCCCA AGCATAGACT GAATGCACAC AGCATCCCAG AGCTATGCTA ACTGAGAGGT   1559
```

CTGGGGTCAG CACTGAAGGC CCACATGAGG AAGACTGATC CTTGGCCATC CTCAGCCCAC 1619

AATGGCAAAT TCTGGATGGT CTAAGAAGGC CCTGGAATTC TAAACTAGAT GATCTGGGCT 1679

CTCTGCACCA TTCATTGTGG CAGTTGGGAC ATTTTTAGGT ATAACAGACA CATACACTTA 1739

GATCAATGCA TCGCTGTACT CCTTGAAATC AGAGCTAGCT TGTTAGAAAA AGAATCAGAG 1799

CCAGGTATAG CGGTGCATGT CATTAATCCC AGCGCTAAAG AGACAGAGAC AGGAGAATCT 1859

CTGTGAGTTC AAGGCCACAT AGAAAGAGCC TGTCTCGGGA GCAGGAAAAA AAAAAAAAAC 1919

GGAATTC 1926

(2) INFORMATION FOR SEQ ID NO:10:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 399 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

Met Ala Met Arg Pro Gly Pro Leu Trp Leu Leu Gly Leu Ala Leu Cys
1 5 10 15

Ala Leu Gly Gly Gly His Gly Pro Arg Pro Pro His Thr Cys Pro Gln
20 25 30

Arg Arg Leu Gly Ala Arg Glu Arg Arg Asp Met Gln Arg Glu Ile Leu
35 40 45

Ala Val Leu Gly Leu Pro Gly Arg Pro Arg Pro Arg Ala Gln Pro Ala
50 55 60

Ala Ala Arg Gln Pro Ala Ser Ala Pro Leu Phe Met Leu Asp Leu Tyr
65 70 75 80

His Ala Met Thr Asp Asp Asp Asp Gly Gly Pro Pro Gln Ala His Leu
85 90 95

Gly Arg Ala Asp Leu Val Met Ser Phe Val Asn Met Val Glu Arg Asp
100 105 110

Arg Thr Leu Gly Tyr Gln Glu Pro His Trp Lys Glu Phe His Phe Asp
115 120 125

Leu Thr Gln Ile Pro Ala Gly Glu Ala Val Thr Ala Ala Glu Phe Arg
130 135 140

Ile Tyr Lys Glu Pro Ser Thr His Pro Leu Asn Thr Thr Leu His Ile
145 150 155 160

Ser Met Phe Glu Val Val Gln Glu His Ser Asn Arg Glu Ser Asp Leu
165 170 175

Phe Phe Leu Asp Leu Gln Thr Leu Arg Ser Gly Asp Glu Gly Trp Leu
180 185 190

Val Leu Asp Ile Thr Ala Ala Ser Asp Arg Trp Leu Leu Asn His His
195 200 205

Lys Asp Leu Gly Leu Arg Leu Tyr Val Glu Thr Ala Asp Gly His Ser
210 215 220

Met Asp Pro Gly Leu Ala Gly Leu Leu Gly Arg Gln Ala Pro Arg Ser
225 230 235 240

Arg Gln Pro Phe Met Val Thr Phe Phe Arg Ala Ser Gln Ser Pro Val
245 250 255

Arg Ala Pro Arg Ala Ala Arg Pro Leu Lys Arg Arg Gln Pro Lys Lys
260 265 270

Thr Asn Glu Leu Pro His Pro Asn Lys Leu Pro Gly Ile Phe Asp Asp
275 280 285

Gly His Gly Ser Arg Gly Arg Glu Val Cys Arg Arg His Glu Leu Tyr
290 295 300

Val Ser Phe Arg Asp Leu Gly Trp Leu Asp Trp Val Ile Ala Pro Gln
305 310 315 320

Gly Tyr Ser Ala Tyr Tyr Cys Glu Gly Glu Cys Ala Phe Pro Leu Asp
325 330 335

Ser Cys Met Asn Ala Thr Asn His Ala Ile Leu Gln Ser Leu Val His
340 345 350

Leu Met Lys Pro Asp Val Val Pro Lys Ala Cys Cys Ala Pro Thr Lys
355 360 365

Leu Ser Ala Thr Ser Val Leu Tyr Tyr Asp Ser Ser Asn Asn Val Ile
370 375 380

Leu Arg Lys His Arg Asn Met Val Val Lys Ala Cys Gly Cys His
385 390 395

(2) INFORMATION FOR SEQ ID NO:11:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 6418 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(ix) FEATURE:
(A) NAME/KEY: misc_feature
(B) LOCATION: 1..6361
(D) OTHER INFORMATION: /note= "HOP-2 genomic sequence"

(ix) FEATURE:
(A) NAME/KEY: exon
(B) LOCATION: 1..837
(D) OTHER INFORMATION: /note= "EXON ONE"

(ix) FEATURE:
(A) NAME/KEY: misc_feature
(B) LOCATION: 884..885
(D) OTHER INFORMATION: /note= "A Gap Occurs Between Positions 884 and 885 in this Sequence"

(ix) FEATURE:
(A) NAME/KEY: exon
(B) LOCATION: 1088..1277
(D) OTHER INFORMATION: /note= "EXON TWO"

(ix) FEATURE:
(A) NAME/KEY: exon
(B) LOCATION: 1350..1814
(D) OTHER INFORMATION: /note= "EXON THREE"

(ix) FEATURE:
(A) NAME/KEY: misc_feature
(B) LOCATION: 1834..1835
(D) OTHER INFORMATION: /note= "A Gap Occurs Between Positions 1834 and 1835 in this Sequence"

(ix) FEATURE:
(A) NAME/KEY: exon
(B) LOCATION: 1883..2077
(D) OTHER INFORMATION: /note= "EXON FOUR"

(ix) FEATURE:
(A) NAME/KEY: exon
(B) LOCATION: 2902..2981
(D) OTHER INFORMATION: /note= "EXON FIVE"

(ix) FEATURE:
(A) NAME/KEY: exon
(B) LOCATION: 3507..3617
(D) OTHER INFORMATION: /note= "EXON SIX"

(ix) FEATURE:
(A) NAME/KEY: exon
(B) LOCATION: 6116..6361
(D) OTHER INFORMATION: /note= "EXON SEVEN"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

GGAATTCCGG CCACAGTGGC GCCGGCAGAG CAGGAGTGGC TGGAGGAGCT GTGGTTGGAG 60

CAGGAGGTGG CACGGCAGGG CTGGAGGGCT CCCTATGAGT GGCGGAGACG GCCCAGGAGG 120

CGCTGGAGCA ACAGCTCCCA CACCGCACCA AGCGGTGGCT GCAGGAGCTC GCCCATCGCC 180

CCTGCGCTGC TCGGACCGCG GCCACAGCCG GACTGGCGGG TACGGCGGCG ACAGACGGAT 240

TGGCCGAGAG TCCCAGTCCG CAGAGTAGCC CCGGCCTCGA GGCGGTGGCG TCCGCGTCCT 300

CTCGTCCAGG AGCCAGGACA GGTGTCGCGC GGCGGGCCGT CCAGGGACCG CGCTGAGGCC 360

GCGGTCGCCC GTCCCGCCCC GCCCCGCCGC CCGCCGCCCG CCGAGCCCAG CCTCCTTGCC 420

GTCGGGGCGT CCCCAGGCCC TGGGTCGGCC GCGGAGCCGA TGCGCGCCCG CTGAGCGCCC 480

CAGCTGAGCG CCCCCGGCCT GCCATGACCG CGCTCCCCGG CCCGCTCTGG CTCCTGGGCC 540

TGGCGCTATG CGCGCTGGGC GGGGGCGGCC CCGGCCTGCG ACCCCCGCCC GGCTGTCCCC 600

AGCGACGTCT GGGCGCGCGC GAGCGCCGGG ACGTGCAGCG CGAGATCCTG GCGGTGCTCG 660

GGCTGCCTGG GCGGCCCCGG CCCCGCGCGC CACCCGCCGC CTCCCGGCTG CCCGCGTCCG 720

CGCCGCTCTT CATGCTGGAC CTGTACCACG CCATGGCCGG CGACGACGAC GAGGACGGCG 780

CGCCCGCGGA GCGGCGCCTG GGCCGCGCCG ACCTGGTCAT GAGCTTCGTT AACATGGGTG 840

AGTGCGGCGC CCGCGCGGGG ACCCTCGGAG TAAACTGGCT GCAGCTGCAG GGCCTCTTCT 900

GGCTCTACAC CCCGGGACCA AGCCTGGAAC AAACGTTTGC ACTAAATGAA GCCGGCCCCA 960

CCCAGGCCTC CCTGGGTCCG CTCCACCTTG AGTGGTGGGT GGCTGGGGGC GGTGGCTCAC 1020

ACCAGCTCTG CCCCCTCCAG AGCCCGAGCC ATTCTGAGTG CCAGCCCAGC GCTGCTTTGT 1080

CTTCTAGTGG AGCGAGACCG TGCCCTGGGC CACCAGGAGC CCCATTGGAA GGAGTTCCGC 1140

TTTGACCTGA CCCAGATCCC GGCTGGGGAG GCGGTCACAG CTGCGGAGTT CCGGATTTAC 1200

AAGGTGCCCA GCATCCACCT GCTCAACAGG ACCCTCCACG TCAGCATGTT CCAGGTGGTC 1260

CAGGAGCAGT CCAACAGGTG CCTTCCCCTT GGCCCGGGTG CCCACCTAAC CCCCCACCTC 1320

ACAGTCTCAT GGTCAAGGCA GCCCAGCAGG GAGTCGTGGT GGGTGAAAGA GAGCCTCAAA 1380

GATGGGAAGG ATGCTTGGCC CGAGGCCCTG CACTGTGGGA AGAGCCCCAG TGACAATCCT 1440

GACTTCAAGT CCCTGCCCTC CATCCTGCTG TGGGGACTTG GACATGGTCA CTGAGACTCA 1500

GTTTCCCCAT GTGTACACCT CTGTGGGCTG AGGCAATGAG ATGAGGCTCA GAAGGGCGCA 1560

GCCAGAGTCA GGTGGGAGAC GCTCCGGTGA CAGCCCCCAG CGGGCCCTGG AGACACGGAG 1620

GCAGCTGTGC CGGCCGCCGG TTAATTGTTC TTTCATGTCC ACAGGGGAGT CTGACTTGTT 1680

CTTTTTGGAT CTTCAGACGC TCCGAGCTGG AGACGAGGGC TGGCTGGTGC TGGATGTCAC 1740

AGCAGCCAGT GACTGCTGGT TGCTGAAGCG TCACAAGGAC CTGGGACTCC GCCTCTATGT 1800

GGAGACTGAG GACGGTGAGG CTGGGGCTCT GCAGCTGCAG AGCCACTGCC CGTGAGTGAC 1860

CCCTCTCTCC TTTCTGTCTC AGGGCACAGC GTGGATCCTG GCCTGGCCGG CCTGCTGGGT 1920

CAACGGGCCC CACGCTCCCA ACAGCCTTTC GTGGTCACTT TCTTCAGGGC CAGTCCGAGT 1980

CCCATCCGCA CCCCTCGGGC AGTGAGGCCA CTGAGGAGGA GGCAGCCGAA GAAAAGCAAC 2040

GAGCTGCCGC AGGCCAACCG ACTCCCAGGG ATCTTTGGTG AGGGTCGGGC AGGCTGGGGC 2100

GAGGCTGTGG CTGTCTGGCT GAGAGAGGCA GGGCGAGAAC CAAGTGGTGG CCCAGAGCCC 2160

AGAGCCTCAG GCTAGGTCGG TTCAAGCTGA CGGCCACTCT CCAGCCACCT TTCCTGACAC 2220

CATCTTGGCC CTGATGCACC CTGGTGACCG GCACTCCGAG GCCTGTCCTG GCTGTCCCTG 2280

CTGCCAGAAG TCTCCCTCTC TCCCCCTGGC TCCTCCGGGT CTTTCTCAGG AGCCTCCTTC 2340

AGAATCAGCT GCCCCTTCCC TGGGAGCCGC AGCCCCTCAT GACCTGCGGT TGTGCCTGGG 2400

CACCTGTGGA TCCTCGGTTG CTTATGCGAT TTTCTCCCCA ACTGGCCAAG CTTCAGGATC 2460

AGGGACAGGC CTGACCCAAC CCCGTGCCCT CCTTCCCAGG GAGTCGGCCC TTGACTGGCC 2520

TGGTCGTGAG CCACTTGAAC CTCGGGAATG GGTGTGGCAG GAGAGGGTGG GCTGGAGTCA 2580

CAGGGGTCTC CAGAGAGGAG GAGGCACAGG ATGGCCGAGG GTCCTGCTGG GCTGTTTACT 2640

GGAGCATAAA GATGCTCATA GGCTGAAGGA CAGGGGAGGA CTGGGCACAG TGTCACTCTA 2700

GCCATTGGGA GCCATGGCAG GCTTCTGAGC TGGGTCATGG TACAAGCAGA GTTCCAGGGA 2760

TGGGCTTTAT GAGCCAAATG GTTTCCTGTC ATTCATTTAT TTGACAAATG TGCTCATCAG 2820

GGCATCCCCC ACCCTGGTAC CCCATAGTAG CTGCACACAG CAGGAACCCC AGAAAAGACC 2880

TTGCCCCTTC TGTCCCTGCA GATGACGTCC ACGGCTCCCA CGGCCGGCAG GTCTGCCGTC 2940

GGCACGAGCT CTACGTCAGC TTCCAGGACC TCGGCTGGCT GGTAATTGCT GACTCTCCTT 3000

GTTTCTGAAA TGACAATCAC CACCTGTAGA TCAGAAGTGA ATCTGCAGGG AGGACATAGA 3060

ATCATGGTGA CTTCAATTTT CTTATGTATT TTTTTCTTCT GTGTTTTCCA AGTTTTCTAA 3120

AGTGAGAATA TGGTGAGAAA GGGTTTTGTT GTTGTTGTTG TGTTTTTTGT TTTTTTTTAA 3180

AAACCCATGA AAATGAAGAC TGAATCAACC AACTAAGCTG TCAGCATTGC CGCAGGGTAA 3240

CTGAGACCTC CCTGCATTGG CTACGACTGC AGCTCTGGGA GGTGTGGGCA GGGGAGGGCC 3300

GGCTGGGGAG GGCCGGCTGG GGAGGGGACA CAAAGTGAAG ATGGGGGTTG TTGGGCCTGA 3360

GCTCCTGCCC AGCCTTTTCC GCCGGGGTTC CTGGGTGGAT TCAAGCCTCT TGGGGGAGAC 3420

GCGCTGCAGG GCTGGAGGAT GGGCTTTGGG CCCTGAGGCT CAGGGAGGAG CACATGGATG 3480

GGACTCACCT TCTCCCTTGC CCCCAGGACT GGGTCATCGC TCCCCAAGGC TACTCGGCCT 3540

ATTACTGTGA GGGGGAGTGC TCCTTCCCAC TGGACTCCTG CATGAATGCC ACCAACCACG 3600

CCATCCTGCA GTCCCTGGTC GATACCGTCG CCCATCCTGC CCAGCCCCCT GGTGGAGGCC 3660

CTGCAGAGAG GGGTCTGGTC CAGCCAGCCG GGAGGCAGTG AGGCCACCTG CTCCATGTCT 3720

CGGGGCTTTG TCTGCACAGA GTCAGTAACG TCGCTAACTT CCCACAGCTC TGCAGGAACT 3780

GGTCCTCATA CAGCCACACT ACTACACATA GACCCACACC CAAACACGGA CACACGTGAA 3840

CAGTCGCGTA TCATGCCTGT TCTATGCACT GAACAAACTC CTGTGGGACA CTTACACACC 3900

TGCGTGCGGC GCTCAGAGGC ACAGCACATG AAACAGATGT GTACACTGTG TGGGGGCTGT 3960

GTGATCTTAA CACACGGGCC CCCGAGTACG CTGGCAAGTC TGACCGCCCG TGATATGTGC 4020

GCACAGTGTG TGGGGTGTGC GTGTGCATCA CCCACCTGTG CCGCACCACA GGTAGGAAGC 4080

TTCTAGATGG TGTGGCTCTC AACCTTTTGG CTTTTTCCCG CAGTTTCTCT CTTGGCTGTC 4140

TGTGTTTTCT CTGGATCCCC TGGCTTTTGA TGCCGTTGGT GTCTGGGGCA ACCTTAAAGG 4200

ACAAAAGCAG GCTTCTGATG GGATCACTGG TGCTGCTCAC CACTGAGTGC TCGTGTGTTT 4260

GCGGATTCTG GCACCGAGGC TTCCTTCTAG AAGTTTTTAC CTAGAATCCC AGTTCCTGGT 4320

ATTGCACAGC CTTATGTTTT CCTCTTAGGA GGTTCAACGG TGATGCCTTG ATCAGGCGCA 4380

GTGGCTCACC CTGTAATCGC AGCACACGAG CCCAGAAGTT CAAGACAAGC CTGAGCAACA 4440

CAGCAAAACC CTGTCTCTAA AATAAAAATT AAAACACACA CACACACACA CACACACACA 4500

CACACACGTG CGCACACAAT GCCTTGGTGT GAGAGGAAAG AAATTACCAA AAGCTGCTCT 4560

GAGCCTATGA TAATACTTCC TTTCTGGGCA GTCAAATGGT GTTTGCTGGA CACCCTGGAG 4620

CCATCTCCTT GGAAAGGCCC AGGGGTGATG AGGAGCTCCG TCGGGGTGGC CTGGCCAGCA 4680

CCTTTATGCC GTGTGGTTCT CACAGCTGCA TGTGTGGGAG GTACATGGGA AGGTGACTGC 4740

ACCTGCGCTC CTGGACTCCA TCTCCTCTGC CCTTGCCCCT GCCCCTCACG TGCAACTAGA 4800

GTGAGTGCTC ACAGCCTACA GGGCAGCAAA CAGGCACTGT GCTCTAGGGG AGGCTGTCGG 4860

TGGGCACAGA AGCAAACCAA CCGTGGAGTT GACACCTCCT GTGAGGAAGA GCAGACGAGC 4920

CGTGCCGTCA GTGGAGTGAG ACTGGGCCCA GCTCTCCACA CAAGGAGGGG CACGTCAGCA 4980

GCTGGAGGAG GAATGTTCCA GAAGGAGCAA GTGCAAGGCC CTAAGACAGG AGCAGGCTGG 5040

CCCTAAGTTC AGGGCAGGGG AGGAGAGGGG CTGGGTGCAG TGAAGGGGAG GAGAGTGGAG 5100

GGAGGTGATC CGGGGTGATA GGCCAGCTCC CGTAGCCTGG GTTCCCTGGG AAGAGGGTGG 5160

ATTTTATTCC AAGCAACCCC AGAGGCTGTC AGAGGTCTTC AGCAAAGAGT GTCCTTGGTC 5220

TGCGTCACCC TCCAGAAGGA CCTTTCTGGC TTGGGGAGGT CGCGGGAGTG GAAGGCAGAG 5280

GAGCAGGGGA TGAGTGAGGG CTGCTGTGGT CACCTGGCAG GTGATGGCAG CTCGACTGGG 5340

CAGGTGGTCC GAGGCAGCAC GGAGGTGGAG GTTGAGCCAG GGGCTGCTCT CAGGGAAGGG 5400

AGGAGGCGAA AGGAGTCATC CAGGAGGCCT CCCAGGCGGG AGCTATGATG TCAGGGCGGG 5460

AGGAATTCTA TGTTCCACTG AGGCCTCATT AGACCCCCAA GTGCAGAAGT GGGAAGGGGA 5520

GCAGGATCCG CAAGTCTGGA GTTCAGAAGA GAGGTCCAAG CTGAGCCAGG GGAGTGGAGA 5580

GGTGCGGGCC AATGCAGGGC CTTGAAGTGC TGAGGGCGGA TCGAGTCCTC TGGGAGAAGG 5640

AGCAGCACAG GAGAGGGGGC GAGGCTGGCT CCCAGAGCCT GGGGAGGGAG GCAGGTGTGG 5700

GGAGGCAGAG CTTGGGGGGG TCTGAAGGGC TATAAGAAGA CAGTGGTCCT TCCAGGTTCC 5760

CCCTTGGACC TCACTAAGGG CACAAACCTG GCCATGAGGT TCTCCTTCCC ATTATCCCCA 5820

GGAGGAAGTC TGAGCCCTTG GCCTGGGACT CGAGGCCCCT CATTAGTGCC CTGCCCACCT 5880

GCCCCACACC CTGGGGCTGC CATGTATCCC TCCCTGGGCA CTGTGGGCAC CACAGCTCCC 5940

GCTCCCAGAG CTCTCAGGGC TGCTCTTATT CCTGTTAATA ATTCTTATTA TTGTGCTGCT 6000

CCCATGTGGC TTGGAGATGG CCAGGGCAGG GAGCAGGTGG AGCTGGGGCG GGCTAGGTGG 6060

GTCCTCAGAG GAGGCCACTG GCTCATGCCC CTGCCTGTGC TCCCTTCCTG GCCAGGTGCA 6120

CCTGATGAAG CCAAACGCAG TCCCCAAGGC GTGCTGTGCA CCCACCAAGC TGAGCGCCAC 6180

CTCTGTGCTC TACTATGACA GCAGCAACAA CGTCATCCTG CGCAAGCACC GCAACATGGT 6240

GGTCAAGGCC TGCGGCTGCC ACTGAGTCAG CCCGCCCAGC CCTACTGCAG CCACCCTTCT 6300

CATCTGGATC GGGCCCTGCA GAGGCAGAAA ACCCTTAAAT GCTGTCACAG CTCAAGCAGG 6360

AGTGTCAGGG GCCCTCACTC TCTGTGCCTA CTTCCTGTCA GGCTTCTGGT CCTTTCTC 6418

(2) INFORMATION FOR SEQ ID NO:12:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 97 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(ix) FEATURE:
(A) NAME/KEY: Protein
(B) LOCATION: 1..97
(D) OTHER INFORMATION: /label= Generic-Seq-7
/note= "wherein each Xaa is independently selected from a group of one or more specified amino acids as defined in the specification."

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

Leu Xaa Xaa Xaa Phe Xaa Xaa Xaa Gly Trp Xaa Xaa Xaa Xaa Xaa Xaa
1 5 10 15

Pro Xaa Xaa Xaa Xaa Ala Xaa Tyr Cys Xaa Gly Xaa Cys Xaa Xaa Pro
20 25 30

Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Asn His Ala Xaa Xaa Xaa Xaa Xaa
35 40 45

Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Cys Cys Xaa Pro
50 55 60

Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Leu Xaa Xaa Xaa Xaa Xaa Xaa Xaa
65 70 75 80

Val Xaa Leu Xaa Xaa Xaa Xaa Xaa Met Xaa Val Xaa Xaa Cys Xaa Cys
85 90 95

Xaa

(2) INFORMATION FOR SEQ ID NO:13:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 102 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(ix) FEATURE:
(A) NAME/KEY: Protein
(B) LOCATION: 1..102
(D) OTHER INFORMATION: /label= Generic-Seq-8
/note= "wherin each Xaa is independently selected from a group of one or more specified amino acids as defined in the specification."

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

Cys Xaa Xaa Xaa Xaa Leu Xaa Xaa Xaa Phe Xaa Xaa Xaa Gly Trp Xaa
1 5 10 15

Xaa Xaa Xaa Xaa Xaa Pro Xaa Xaa Xaa Xaa Ala Xaa Tyr Cys Xaa Gly
20 25 30

Xaa Cys Xaa Xaa Pro Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Asn His Ala
35 40 45

Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa
50 55 60

Xaa Cys Cys Xaa Pro Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Leu Xaa Xaa
65 70 75 80

Xaa Xaa Xaa Xaa Xaa Val Xaa Leu Xaa Xaa Xaa Xaa Xaa Met Xaa Val
85 90 95

Xaa Xaa Cys Xaa Cys Xaa
100

What is claimed is:

1. A substantially pure protein comprising the amino acid sequence defined by residues 303 to 399 of Seq. ID No. 1, including allelic, species and other amino acid sequence variants thereof.

2. The protein of claim 1 wherein said amino acid sequence is defined by residues 298 to 399 of Seq. ID No. 1, including allelic, species and other amino acid sequence variants thereof.

3. The protein of claim 2 wherein said amino acid sequence is defined by residues 264 to 399 of Seq. ID No. 1, including allelic, species and other amino acid sequence variants thereof.

4. The protein of claim 3 wherein said amino acid sequence is defined by residues 261 to 399 of Seq. ID No. 1, including allelic, species and other amino acid sequence variants thereof.

5. The protein of claim 4 wherein said amino acid sequence is defined by residues 18 to 399 of Seq. ID No. 1, including allelic, species and other amino acid sequence variants thereof.

6. The protein of claim 5 wherein said amino acid sequence is defined by residues 1 to 399 of Seq. ID No. 1, including allelic, species and other amino acid sequence variants thereof.

7. A substantially pure antibody which binds to an epitope on a protein encoded by a nucleic acid comprising the DNA sequence defined by bases 69-1265 of Seq. ID No. 1.

8. A substantially pure nucleic acid comprising part or all of the DNA sequence defined by bases 1 to 1674 of Seq. ID No. 1, including allelic, species and other amino acid sequence variants thereof.

9. A substantially pure nucleic acid encoding a morphogenically active protein, said nucleic acid comprising a DNA sequence that hybridizes to part or all of the DNA sequence defined by bases 120 to 848 of Seq. ID No. 1, under stringent conditions, including allelic, species and other amino acid sequence variants thereof.

10. A substantially pure nucleic acid comprising part or all of the DNA sequence defined by bases 120 to 848 of Seq. ID No. 1, including allelic, species and other amino acid sequence variants thereof.

11. A substantially pure nucleic acid encoding a morphogenic protein, said nucleic acid comprising a DNA sequence defined by bases 975 to 1265 of Seq. ID No. 1, including allelic, species and other amino acid sequence variants thereof.

12. The nucleic acid of claim 11 wherein said DNA sequence is defined by bases 960 to 1265 of Seq. ID No. 1, including allelic, species and other amino acid sequence variants thereof.

13. The nucleic acid of claim 12 wherein said DNA sequence is defined by bases 858 to 1265 of Seq. ID No. 1, including allelic, species and other amino acid sequence variants thereof.

14. The nucleic acid of claim 13 wherein said DNA sequence is defined by bases 849 to 1265 of Seq. ID No. 1, including allelic, species and other amino acid sequence variants thereof.

15. The nucleic acid of claim 14 wherein said DNA sequence is defined by bases 120 to 1265 of Seq. ID No. 1, including allelic, species and other amino acid sequence variants thereof.

16. The nucleic acid of claim 15 wherein said DNA sequence is defined by bases 69-1265 of Seq. ID No. 1, including allelic, species and other amino acid sequence variants thereof.

17. A vector comprising at least part of the nucleic acid sequence defined by bases 69-1265 of Seq. ID No. 1, including allelic, species and other amino acid sequence variants thereof, wherein said sequence is sufficient to encode a morphogenic protein.

18. A cell transformed with the vector of claim 17.

19. A cell adapted to express a nucleic acid comprising the sequence defined by bases 975-1265 of Seq. ID No. 1, including allelic, species or amino acid sequence variants thereof.

20. A substantially pure morphogenic protein encoded by at least part of the nucleic acid sequence of Seq. ID No. 1 including allelic, species and other amino acid sequence variants thereof.

21. A composition for increasing the progenitor cell population in a mammal comprising:
progenitor cells, stimulated ex vivo by exposure to a morphogenically active fragment of OP-3, or allelic or species variants thereof, at a concentration and for a time sufficient such that said progenitor cells are stimulated to proliferate.

22. A composition for inducing tissue growth in a mammal comprising:
progenitor cells, stimulated by exposure to a morphogenically active fragment of OP-3, or allelic or species variants thereof, at a concentration and for a time sufficient such that said progenitor cells, when disposed in vivo within a tissue locus, are capable of tissue-specific differentiation and proliferation within said locus.

23. The composition of claim 21 or 22 wherein said morphogenically active fragment comprises amino acid residues 303 to 399 of Seq. ID No. 1 or allelic, species and other variants thereof.

24. The composition of claim 23 wherein said morphogenically active fragment comprises amino acid residues 298 to 399 of Seq. ID No. 1 or allelic, species and other variants thereof.

25. The composition of claim 24 wherein said morphogenically active fragment comprises amino acid residues 264 to 399 of Seq. ID No. 1 or allelic, species and other variants thereof.

26. The composition of claim 25 wherein said morphogenically active fragment comprises amino acid residues 261 to 399 of Seq. ID No. 1 or allelic, species and other variants thereof.

27. The composition of claim 26 wherein said morphogenically active fragment comprises amino acid residues 18 to 399 of Seq. ID No. 1 or allelic, or species and other amino acid sequence variants thereof.

28. The composition of claim 27 wherein said morphogenically active fragment comprises amino acid residues 1 to 399 of Seq. ID No. 1 or allelic, species and other amino acid sequence variants thereof.

29. The composition of claim 21 or 22 wherein said progenitor cells are hemopoietic pluripotential stem cells.

30. The composition of claim 21 or 22 wherein said progenitor cells are of mesenchymal origin.

31. A composition for inducing the formation of replacement tissue at a tissue locus in a mammal comprising:
a biocompatible, acellular matrix having components specific for said tissue and capable of providing a morphogenically permissive, tissue-specific environment; and
a morphogenically active fragment of OP-3, or allelic or species variants thereof, for inducing the developmental cascade of tissue morphogenesis at said locus when absorbed on said matrix and provided to a tissue-specific locus requiring replacement tissue.

32. A composition for inducing the formation of replacement tissue at a tissue locus in a mammal comprising:
a biocompatible, acellular matrix capable of providing a morphogenically permissive environment; and a morphogenically active fragment of OP-3, or allelic or species variants thereof, for inducing the developmental cascade of tissue morphogenesis at said locus when absorbed on said matrix and provided to a tissue-specific locus requiring replacement tissue.

33. The composition of claim 31 or 32 wherein said morphogenically active fragment comprises amino acid residues 303 to 399 of Seq. ID No. 1, including allelic, species and other amino acid sequence variants thereof.

34. The composition of claim 33 wherein said morphogenically active fragment comprises amino acid residues 298 to 399 of Seq. ID No. 1, including allelic, species and other amino acid sequence variants thereof.

35. The composition of claim 34 wherein said morphogenically active fragment comprises amino acid residues 264 to 399 of Seq. ID No. 1, including allelic, species and other amino acid sequence variants thereof.

36. The composition of claim 35 wherein said morphogenically active fragment comprises amino acid residues 261 to 399 of Seq. ID No. 1, including allelic, species and other amino acid sequence variants thereof.

37. The composition of claim 36 wherein said morphogenically active fragment comprises amino acid residues 18 to 399 of Seq. ID No. 1, including allelic, species and other amino acid sequence variants thereof.

38. The composition of claim 32 wherein said morphogenically active fragment comprises amino acid residues 1 to 399 of Seq. ID No. 1, including allelic, species and other amino acid sequence variants thereof.

39. The composition of claim 31 or 32 wherein said matrix is biodegradable.

40. The composition of claim 31 or 32 wherein said matrix is derived from organ-specific tissue.

41. The composition of claim 31 or 32 wherein said matrix comprises collagen and cell attachment factors selected from the group consisting of glycosaminoglycans and proteoglycans.

42. The composition of claim 31 or 32 wherein said matrix defines a structure which permits the attachment, differentiation and proliferation of migratory progenitor cells from the body of said mammal.

43. A method of increasing a population of progenitor cells comprising the step of: contacting progenitor cells with a morphogenically active fragment of OP-3, or allelic or species variants thereof, at a concentration and for a time sufficient such that said progenitor cells are stimulated to proliferate.

44. The method of claim 43 further comprising the step of:
supplying said stimulated progenitor cells to a mammal to increase the progenitor cell population in said mammal.

45. A method of inducing tissue growth in a mammal comprising the step of:
contacting progenitor cells with a morphogenically active fragment of OP-3, or allelic or species variants thereof, at a concentration and for a time sufficient such that said progenitor cells, when provided to a tissue-specific locus in a mammal, are capable of tissue-specific differentiation and proliferation at said locus.

46. The method of claim 43 or 45 wherein said progenitor cells are of mesenchymal origin.

47. A method of maintaining the phenotypic expression of differentiated cells in a mammal comprising the steps of:
contacting said differentiated cells with a morphogenically active fragment of OP-3, or allelic or species variants thereof, at a concentration and for a time sufficient such that said cells are stimulated to express their phenotype.

48. The method of claim 47 wherein said differentiated cells are senescent or quiescent cells.

49. A method of inducing tissue growth at a tissue locus in a mammal comprising:
providing said locus with a morphogenically active fragment of OP-3, or allelic or species variants thereof, at a concentration and for a time sufficient such that said morphogenically active fragment, when provided to a morphogenically permissive tissue-specific locus, is capable of inducing the developmental cascade of tissue morphogenesis at said locus.

50. The method of claim 49 wherein said tissue is hepatic tissue, and said tissue locus is the liver.

51. The method of claim 49 wherein said tissue is cartilage or bone tissue, and said tissue locus is osteoporotic bone.

52. The method of claim 49 wherein said OP-3, or allelic or species variants thereof, is provided to said locus in association with a biocompatible, acellular matrix.

53. The method of claim 52 wherein said matrix has components specific for said tissue.

54. The method of claim 52 wherein said matrix is biodegradable.

55. The method of claim 52 wherein said matrix is derived from organ-specific tissue.

56. The method of claim 52 wherein said matrix comprises collagen and cell attachment factors specific for said tissue.

57. The method of claim 52 wherein said matrix defines a structure which permits the attachment, differentiation and proliferation of migratory progenitor cells from the body of said mammal.

58. The method of claims 43, 45, 47 or 49 wherein said morphogenically active fragment comprises amino acid residues 303 to 399 of Seq. ID No. 1, including allelic, species and other amino acid sequence variants thereof.

59. The method of claim 58 wherein said morphogenically active fragment comprises amino acid residues 298 to 399 of Seq. ID No. 1, including allelic, species and other amino acid sequence variants thereof.

60. The method of claim 59 wherein said morphogenically active fragment comprises amino acid residues 264 to 399 of Seq. ID No. 1, including allelic, species and other amino acid sequence variants thereof.

61. The method of claim 60 wherein said morphogenically active fragment comprises amino acid residues 261 to 399 of Seq. ID No. 1, including allelic, species and other amino acid sequence variants thereof.

62. The method of claim 61 wherein said morphogenically active fragment comprises amino acid residues 18 to 399 of Seq. ID No. 1, including allelic, species and other amino acid sequence variants thereof.

63. The method of claim 62 wherein said morphogenically active fragment comprises amino acid residues 1 to 399 of Seq. ID No. 1, including allelic, species and other amino acid sequence variants thereof.

64. A method of producing a morphogenically active protein comprising the steps of:
transfecting cells with a nucleic acid sequence encoding a morphogenic protein comprising the amino acid residues 303 to 399 of Seq. ID No. 1;
culturing said cells in a suitable culture medium;
expressing said morphogenic protein from said nucleic acid; and
isolating and purifying said protein from said culture medium.

65. The method of claim 64 wherein said morphogenic protein comprises amino acid residues 298 to 399 of Seq. ID No. 1, including allelic, species and other amino acid sequence variants thereof.

66. The method of claim 65 wherein said morphogenic protein comprises amino acid residues 264 to 399 of Seq. ID No. 1, including allelic, species and other amino acid sequence variants thereof.

67. The method of claim 66 wherein said morphogenic protein comprises amino acid residues 261 to 399 of Seq. ID No. 1, including allelic, species and other amino acid sequence variants thereof.

68. The method of claim 67 wherein said morphogenic protein comprises amino acid residues 18 to 399 of Seq. ID No. 1, including allelic, species and other amino acid sequence variants thereof.

69. The method of claim 68 wherein said morphogenic protein comprises amino acid residues 1 to 399 of Seq. ID No. 1, including allelic, species and other amino acid sequence variants thereof.

70. The protein of claim 1, 2 or 3 wherein said amino acid sequence variant has an amino acid substitution for the serine at position 315 or the cysteine at position 338 in Seq. ID No. 1.

71. The protein of claim 70 wherein said amino acid sequence variant has a tryptophan residue in place of the serine at position 315 in Seq. ID No. 1.

72. The protein of claim 70 wherein said cysteine residue at position 338 in Seq. ID No. 1 is substituted for an amino acid selected from the group consisting of tyrosine, histidine, isoleucine and serine.

73. A chimeric morphogen comprising the amino acid sequence of claim 1, 2 or 3.

74. The composition of claim 22, 23, 31 or 32 wherein said morphogen amino acid sequence variant has an amino acid substitution for the serine at position 315 or the cysteine at position 338 in Seq. ID No. 1.

75. The method of claim 43, 45, 47 or 49 wherein said morphogen amino acid sequence variant has an amino acid substitution for the serine at position 315 or the cysteine at position 338 in Seq. ID No. 1.

76. A morphogen comprising an amino acid sequence defined by Generic Sequence 7 or 8 (Seq. ID Nos. 12 or 13.)

77. The protein of claims 1, 2, 3 or 4 wherein said protein comprises a dimeric protein species complexed with a peptide comprising a pro region of a member of the morphogen family, or an allelic, species or other sequence variant thereof.

78. The protein of claim 77 wherein said dimeric protein species is noncovalently complexed with said peptide.

79. The protein of claim 77 wherein said dimeric protein species is complexed with two said peptides.

80. The protein of claim 77 wherein said peptide comprises at least the first 18 amino acids of a sequence defining said pro region.

81. The protein of claim 80 wherein said peptide comprises the full length form of said pro region.

82. The protein of claim 77 wherein said peptide comprises a nucleic acid that hybridizes under stringent conditions with a DNA defined by nucleotides 136-192 of Seq. ID No. 3 or nucleotides 157-211 of Seq. ID No. 7.

83. the protein of claim 77 wherein said peptide comprises at least the first 18 amino acid of the pro region of OP3 (Seq. ID No. 1).

84. The protein of claim 77 wherein said complex is further stabilized by exposure to a basic amino acid, a detergent or a carrier protein.

Fig. 1.1

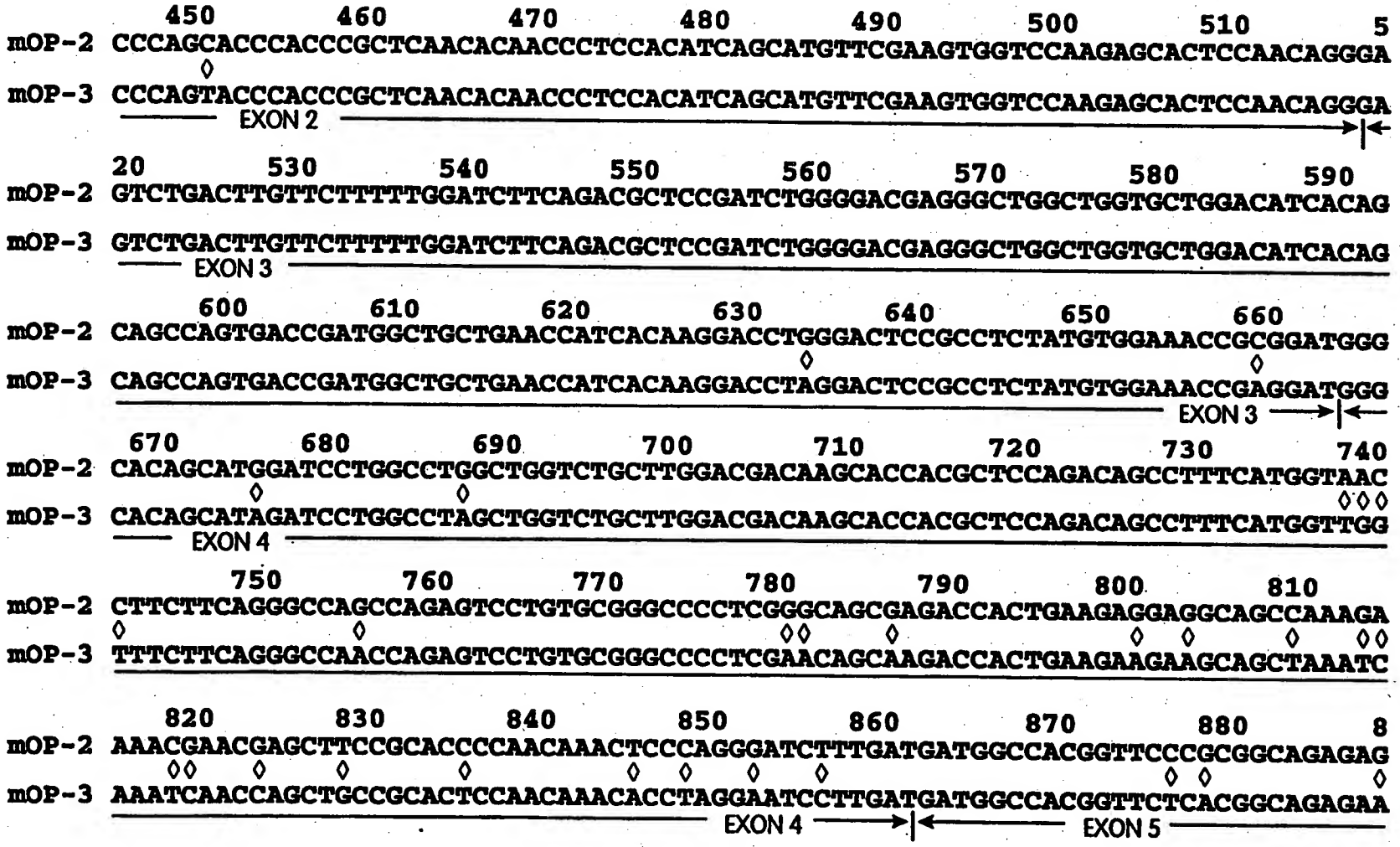

Fig. 1.2

```
          90        900       910       920       930       940       950       960
mOP-2 GTTTGCCGCAGGCATGAGCTCTACGTCAGCTTCCGTGACCTTGGCTGGACTGGGTCATCGCCCCCCAGGG
                                ◇                                   ◇◇      ◇
mOP-3 GTTTGCCGCAGGCATGAGCTCTATGTCAGCTTCCGTGACCTTGGCTGGACTCTGTCATTGCCCCCCAGGG
      ──────────────────────────────────────────────── EXON 5 ──►┼◄── EXON 6 ──────

          970       980       990      1000      1010      1020      1030
mOP-2 CTACTCTGCCTATTACTGTGAGGGGGAGTGTGCTTTCCCACTGGACTCCTGTATGAACGCCACCAACCATGCCA
            ◇         ◇◇        ◇◇◇ ◇       ◇         ◇                   ◇
mOP-3 CTACTCCGCCTATTACTGTGCTGGGGAGTGCATCTACCCACTGAACTCCTGTATGAACTCCACCAACCACGCCA
      ──────────────────────────────────────────────────────────────────────────

         1040      1050      1060      1070      1080      1090      1100      1110
mOP-2 TCTTGCAGTCTCTGGTGCACCTGATGAAGCCAGATGTTGTCCCCAAGGCATGCTGTGCACCCACCAAACTGAGT
      ◇◇◇       ◇ ◇       ◇         ◇ ◇◇                ◇◇        ◇   ◇◇  ◇
mOP-3 CTATGCAGGCCCTGGTACATCTGATGAAGCCAGATATCATCCCCAAGGTGTGCTGTGGCCTACTGAGCTGAGT
      ──────── EXON 6 ──►┼◄── EXON 7 ──────────────────────────────────────────

         1120      1130      1140      1150      1160      1170      1180
mOP-2 GCCACCTCTGTGCTGTACTATGACAGCCAGCAACAATGTCATCCTGCGTAAACACCGTAACATGGTGGTCAAGGC
      ◇◇        ◇         ◇ ◇       ◇         ◇ ◇◇◇ ◇     ◇         ◇
mOP-3 GCCATTTCTCTGCTCTACTATGATAGAACAATAATGTCATCCTGCGCAGGGAGCGCAACATGGTAGTCCAGGC
      ──────────────────────────────────────────────────────────────────────────

         1190      1200
mOP-2 CTGTGGCTGCCACTGA
mOP-3 CTGTGGCTGCCACTGA
      ──────── EXON 7 ──►
```

Fig. 1.3

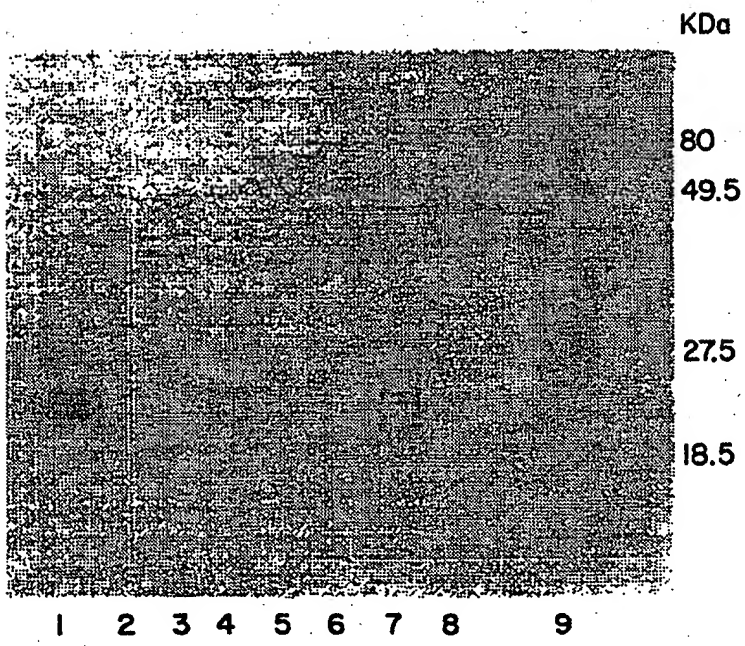

Fig. 2